Evaluation of the Arctic surface radiation budget in CMIP5 models

Robyn C. Boeke

Science Systems Applications Inc., Hampton, Virginia

Patrick C. Taylor

NASA Langley Research Center, Climate Science Branch, Hampton, Virginia

Key Points:

1. Biases in the Arctic surface radiation budget from CMIP3 remain in CMIP5
2. The surface warming effect of cloud is too small in models in winter
3. Significant spatial variations are found in the Arctic surface radiative flux biases

Corresponding author address: Robyn C. Boeke, Science Systems Applications, Inc., 1 Enterprise Parkway, Suite 200, Hampton, VA, 23666 (robyn.c.boeke@nasa.gov).

Abstract

The Arctic region is warming at a rate more than double the global average. This trend is predicted to continue in the coming decades as simulated in the Coupled Model Intercomparison Project 5 (CMIP5) climate projections. Despite the consistency in the projected Arctic warming rate relative to the globe, significant inter-model spread is found in the simulated present-day Arctic surface temperature and the warming response to inCREased CO_2. The representation of the surface radiation budget is a key factor in the simulation of the Arctic climate system. The goal of this paper is to evaluate the representation of the Arctic surface radiation budget in CMIP5 and investigate the influence of these biases on the simulated present-day Arctic climate. First, the Arctic surface radiation budget CMIP5 Historical forcing scenario for 17 models is evaluated against Cloud and Earth's Radiant Energy System (CERES) surface radiative fluxes for the annual mean and seasonal cycle. The CMIP5 multi-model ensemble is found to simulate longwave surface fluxes well during the sunlit months but produce wintertime values that are less than observations. Shortwave fluxes show substantial across-model spread during summer. Second, the independent column approximation (ICA) is used to attribute the radiative flux biases to clear-sky, cloudy-sky, and cloud fraction contributions. Lastly, the surface radiation budget biases are used to explain the biases in simulated present day Arctic surface temperature.

1. Introduction

Arctic surface temperature is increasing at a rate outpacing globally-averaged warming by 2-3 times over the last 50 years [Chylek et al., 2009; ACIA, 2005]. Amplified warming of the Arctic surface—referred to as polar or Arctic amplification—is a robust climate system response to an external forcing [e.g., Manabe and Wetherald, 1975; Hansen et al., 1984; Rind, 1987; Lu and Cai, 2012; Taylor et al., 2013; Sejas et al., 2014]. Despite the fact that Arctic amplification is a robust feature of Coupled Model Intercomparison Project (CMIP) CMIP3 (IPCC 2007) and CMIP5 climate models (IPCC 2013), the largest intermodel spread in surface temperature warming is found in the Arctic [e.g., Pithan and Mauritsen, 2014]. The magnitude of Arctic surface temperature warming has significant implications for the projected changes in other features of the Arctic climate system including sea ice extent, land ice sheet mass, and clouds. Reduction in the intermodel spread in Arctic surface temperature projections is imperative not just to improve Arctic climate projections but also for projections in the global climate system.

Amplified warming of the Arctic surface is attributed to a number of radiative and nonradiative feedback processes. These processes include surface albedo feedback [e.g., Manabe and Wetherald, 1975; Hall, 2004], atmospheric and ocean dynamical transport feedbacks [e.g., Holland and Bitz, 2003; Cai, 2005; Cai, 2006; Graversen and Wang, 2009; Bitz et al., 2012; Langen et al., 2012], and cloud feedbacks [e.g., Holland and Bitz, 2003; Vavrus, 2004; Lu and Cai, 2009; Taylor et al., 2011a,b; Taylor et al., 2013]. Each process affects the Arctic surface temperature through perturbations to the Arctic surface radiation budget. Therefore, evaluating and understanding the physical causes of biases within the simulated Arctic surface radiation budget is important for constraining climate model projections.

The present day observed and model simulated Arctic radiation budget has been examined in many studies [Kay and L'Ecuyer, 2013; Karlsson and Svensson, 2011; English et al., 2015]. An 11-year cloud and radiation climatology was presented in Kay and L'Ecuyer [2013] by combining observational datasets including CloudSat, CALIPSO, MODIS, and CERES-EBAF and found that the uncertainty in flux calculations is primarily a result of cloud uncertainty. On average, oceanic clouds warm the Arctic surface 10 W m^{-2} annually although with large spatial, seasonal, and interannual variability in the cloud forcing. Karlsson and Svensson [2011] presented a methodology of studying Arctic clouds and surface fluxes using the CMIP3 multi-model dataset and observations from the AVHRR Polar Pathfinder (APP-x) product. The findings indicate that GCMs have difficulty simulating the surface cloud radiative effect (*CRE*), particularly in winter. This difference is attributed to the large intermodel spread in wintertime cloud fraction—ranging from 36 to 94%. Karlsson and Svensson [2011] found no clear relation between the model spread in wintertime cloud forcing and surface temperature, indicating that other processes are responsible for the intermodel spread in Arctic surface temperature.

This study provides an extensive evaluation of the Arctic surface radiation budget as simulated by GCMs and updates previous studies by using the CMIP5 models [Taylor et al., 2012]. Two high-quality data sets of Arctic surface radiation fluxes are used: CloudSAT-CALIPSO-CERES-MODIS (C3M) [Kato et al., 2010; Kato et al., 2011a] and Clouds and Earth's Radiant Energy System (CERES) Surface Energy Balanced and Filled (SFC-EBAF) v2.8 [Kato et al., 2011b], as well as CMIP5 model simulations from the Historical scenario. The data set and model simulations are discussed in Section 2. An annual mean and seasonal cycle evaluation of the Arctic surface radiation budget is presented in Section 3. Section 4 attributes the surface radiative budget biases between clear and cloudy sky fluxes and total column cloud fraction

using the independent column approximation. A broader discussion linking model skill in simulating Arctic surface energy balance and surface temperature is presented in Section 5.

2. Data Sets and Models

a. Surface Radiative Fluxes and Cloud Fraction Data Sets

CERES SFC-EBAF v2.8 [Kato et al., 2013] provides gridded, monthly mean surface radiative fluxes from March 2000 through present. Monthly average radiative fluxes are provided on a 1° equal-area grid. The CERES SFC-EBAF data set is used as the primary source of surface radiative flux terms including upwelling and downwelling shortwave and longwave radiation for all-sky and clear-sky.

The CERES SFC-EBAF data set provides the only surface radiative flux data set across the Arctic constrained by and consistent with observed top-of-atmosphere (TOA) fluxes (CERES TOA-EBAF) [Loeb et al., 2009]. The CERES SFC-EBAF radiative fluxes are determined by first using satellite-retrieved surface, cloud, and aerosol information and temperature and humidity profiles from the Goddard Earth Observing System (GEOS-4 and 5) Data Assimilation System [Bloom et al., 2005; Rienecker et al., 2008] as input into a radiative transfer model. A Lagrange multiplier scheme is then used to objectively adjust the radiative transfer model inputs based upon uncertainty estimates such that the computed TOA fluxes are consistent with CERES observations. The CERES SFC-EBAF fluxes are computed using the adjusted inputs. Kato et al., [2013] state that monthly mean root mean square differences between CERES SFC-EBAF and surface observations are 13.3 W m^{-2} and 7.8 W m^{-2} in the downwelling shortwave and 7.1 W m^{-2} and 7.8 W m^{-2} for the downwelling longwave over ocean and land, respectively. Kato et al., [2013] provide an in depth discussion of the data set methodology and evaluation.

The Ed RalB1 CALIPSO-CloudSat-CERES-MODIS (C3M) data set provides cloud property and radiative flux data from July 2006 through June 2010 [Kato et al., 2010; Kato et al., 2011]. C3M is a merged data set collocating Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) [Winker et al., 2010], CloudSat Cloud Profiling Radar (CPR) [Stephens et al., 2008], CERES [Wielicki et al. 1996], and Aqua Moderate Resolution Imaging Spectrometer (MODIS) onto the CERES 20-km footprint. C3M cloud property vertical profiles are constructed from CALIPSO version 3 Vertical Feature Mask (VFM; 30 m vertical resolution below 8.2 km) [Winker et al., 2007] and the CloudSAT release 04 CLDCLASS product (240 m vertical resolution) [Sassen and Wang, 2008]. The C3M vertical profile merges cloud top and base heights determined separately by the VFM and CLDCLASS products, however preference is given to the CALIPSO VFM cloud boundaries. CPR-derived cloud boundaries are used only when CALIOP does not detect a cloud or is completely attenuated [Kato et al., 2010].

Clear- and all-sky surface radiative fluxes are provided within C3M. C3M surface radiative fluxes are computed using GEOS-5 temperature and humidity profiles and the merged cloud vertical profiles and a modified Fu-Liou radiative transfer model [Fu and Liou, 1993; Fu et al., 1997; Kratz and Rose, 1999; Kato et al., 1999, 2005]. Kato et al., [2011; 2013] indicate significant improvements in computed TOA and SFC fluxes when using CALIPSO and CloudSat-derived cloud properties.

b. Climate Models

The Coupled Model Intercomparison Project Phase 5 (CMIP5) is a collaboration of 20 international climate modeling groups to conduct a coordinated set of climate model experiments [Taylor et al., 2012]. The Historical simulation is an ocean-atmosphere coupled model forced with the best-estimate natural and anthropogenic radiative forcing from 1850 through 2005.

Output from 17 models running this simulation is used (summarized in Table 1) and accessed from http://pcmdi.llnl.gov/. The models used in this study were chosen based on available model output: surface and TOA fluxes, total column cloud fraction, and surface temperature.

3. Evaluation of Surface Radiative Fluxes in the Arctic

a. Annual mean

The annual mean surface radiative flux differences between CMIP5 models and CERES SFC-EBAF are quantified in the Arctic domain average and spatially. Table 2 summarizes the Arctic domain-averaged—defined as grid points poleward of 66°N—observed and modeled surface radiative fluxes: all-sky downwelling longwave radiation (RLDS), clear-sky downwelling longwave radiation (RLDSCS), all-sky downwelling shortwave radiation (RSDS), clear-sky downwelling shortwave radiation (RSDSCS), all-sky upwelling shortwave radiation (RSUS), and clear-sky upwelling shortwave radiation (RSUSCS). Figure 1 shows the spatial distribution of the annual mean Arctic surface radiative fluxes for CERES SFC-EBAF (first column), the difference plot of the multi-model ensemble mean and CERES SFC-EBAF (second column), and the across-model standard deviation (third column). The standard deviation represents the level of agreement between models.

The spatial pattern of the annual mean RLDS (Fig. 1a) largely follows the distribution of cloud cover and surface air temperature. The highest values are in the North Atlantic, where cloud cover is extensive and thick and Arctic air temperatures and specific humidity are highest [Serreze and Barry 2014]. The lowest RLDS values are found over Greenland since the high elevation is associated with colder air temperatures and lower water vapor amounts. The ensemble mean underestimates CERES RLDS by 9.95 W m^{-2} in the Arctic domain average (Table 2). The negative differences are largest in the North Atlantic and Baffin Bay—exceeding

-20 W m^{-2}. The model standard deviation (Fig. 1c) reveals that the largest intermodel spread in RLDS is also found in the North Atlantic region. The large model standard deviations of RLDS indicate important intermodel differences in the quantities that control these surface fluxes: atmospheric temperature, humidity, and cloud properties.

The characteristics of the RLDSCS differences between the models and CERES SFC-EBAF are similar to those of RLDS. The ensemble mean – CERES-SFC EBAF RLDSCS (Fig. 1e) underestimates CERES by 6.19 W m^{-2} in the domain-wide annual mean (Table 2). Additionally, the North Atlantic contributes most to the domain-wide bias that also exceeds -20 W m^{-2}. The largest intermodel RLDSCS standard deviations are also in the North Atlantic, indicating significant differences in simulated atmospheric temperature and humidity.

The sign of the RSDS differences between CERES SFC-EBAF and the ensemble (Fig. 1h) oppose those of RLDS. The largest positive differences are found over the North Atlantic. As with the RLDS differences, this is consistent with the simulation of clouds that are either too few or too thin. The largest negative RSDS differences are found over Greenland and Northern coastal Alaska exceeding 10 W m^{-2}. An Arctic domain-averaged RSDS ensemble minus CERES difference of 2.92 W m^{-2} is found due to the offset of too much RSDS over the North Atlantic and too little RSDS over the rest of the Arctic. Model standard deviation in RSDS (Fig. 1i) is weaker than RLDS and shows the most disagreement over land. However, separating the results by surface type, the ensemble mean captures observed RSDS better over land than ocean.

The RSDSCS differences between the models and observations (Fig. 1k) indicate a bias of the opposite sign of RSDS. The domain-averaged ensemble mean RSDSCS is 4.86 W m^{-2} greater than CERES. Models show too much RSDSCS over the central Arctic (approaching a 5 W m^{-2} bias) and North Atlantic region, where biases can exceed 15 W m^{-2}. In contrast to the model and

CERES RSDS differences, all of the Arctic, except Greenland, exhibits the same sign of RSDSCS differences. The model standard deviation (Fig. 1l) is low over the entire domain.

Model and CERES differences for RSUS and RSUSCS are shown in Figs. 1m-r. These clear-sky and all-sky differences are similar to each other and illustrate the intermodel differences in surface albedo. While model differences with CERES in RSUS and RSUSCS are of the same magnitude of other flux terms, model standard deviation is greater than for any other surface flux and specifically in regions where sea ice is found in observations.

Alluded to above, the radiative effects of clouds make a large contribution to the Arctic surface radiation budget and influence the difference between models and observations. To quantify the impact of clouds on the surface radiation budget, the cloud radiative effect (CRE) is calculated. CRE is defined for surface shortwave (SW), longwave (LW), and net fluxes as the all-sky minus clear-sky flux difference, where a downward flux is positive. Surface albedo (α) is calculated using shortwave clear-sky fluxes. LW, SW, Net CRE and α are defined as

$$LW\ CRE = (RLDS - RLDSCS) \quad (1)$$

$$SW\ CRE = (RSDS - RSDSCS) \cdot (1 - \alpha) \quad (2)$$

$$NET\ CRE = (RSDS - RSDSCS) \cdot (1 - \alpha) + (RLDS - RLDSCS) \quad (3)$$

$$\alpha = \frac{RSUSCS}{RSDSCS} \quad (4)$$

SW CRE is usually negative because RSDS decreases in the presence of clouds as clouds reflect more solar radiation to space. The presence of α in eq. (2) indicates that in the absence of a change in clouds, a change in α will change SW CRE. LW CRE is typically positive because clouds enhance the emitted downwelling longwave radiation to the surface. The Net CRE is simply the result of adding the LW and SW CRE. Figure 2 shows the annual mean spatial distribution of the LW, SW, and Net CRE over the Arctic domain. Table 3 summarizes the

corresponding annual, domain-wide LW, SW, and Net CRE mean values for models and CERES SFC-EBAF observations.

Annually, the ensemble mean underestimates CERES LW CRE (Table 3) by 3.60 W m^{-2}. Figure 2b illustrates that the negative differences are found over most of the Arctic. The bias is largest over the North Atlantic and Greenland where the ensemble mean LW CRE values are 12 W m^{-2} lower than CERES SFC-EBAF. These differences indicate that model clouds are either too few or too thin; models have difficulty capturing the insulating effect of clouds, leading to an underestimate in LW CRE. This result is consistent with results from CMIP3 [Sorteberg et al., 2007; Karlsson and Svensson, 2011]. The North Atlantic storm track and Greenland regions show the largest intermodel standard deviations in LW CRE. Storm track dynamics is an important mechanism for cloud generation in the North Atlantic; thus, it is likely that model cloud differences in this region are influenced by differences in the storm track.

SW CRE is negative throughout the sun lit portion of the year (Mar-Oct) and produces a cooling effect. It is most negative in the North Atlantic where clouds are prevalent and least negative over Greenland due to the high surface albedo. Annual-average ensemble SW CRE is 3.76 W m^{-2} more negative than CERES, and model standard deviation is large, especially over land and the North Atlantic. Unlike other variables, a moderate spatial correlation is found between the ensemble mean - CERES SW CRE difference and the standard deviation of the models ($r = 0.63$). This correlation indicates that regions where the models disagree are also regions with larger biases. One interpretation of this correlation is that the models disagree about the processes that control the SW CRE (indicated by the model standard deviation) and that the collection of models does not capture processes correctly (indicated by large biases).

The Arctic domain-averaged ensemble mean Net CRE is biased low by 7.36 W m^{-2} in the annual average. The sign of the bias is consistent across the entire domain, with larger biases over land than over the central Arctic Ocean. The model standard deviation (Fig. 2i) illustrates the difficulty that all models have in simulating the Net CRE over land and the North Atlantic.

b. Seasonal Cycle

The Arctic domain-averaged seasonal cycle of the six surface fluxes is calculated and shown in Fig. 3. CERES SFC-EBAF observations are plotted in black. The grey region is the 90% confidence interval for the difference in means between the ensemble and CERES: if the shaded region contains the CERES points, then it indicates that the ensemble mean is in agreement with the observed fluxes. The seasonal cycle for the CMIP5 models is calculated using a 5-year period from 2000 through 2005 in order to overlap with the anthropogenic forcing observed by CERES.

Models simulate too little wintertime RLDS and RLDSCS, whereas in summer models show good agreement with observations. The low bias in RLDS is a long-standing problem with climate models [e.g., Karlsson and Svensson 2011]. The presence of lower RLDSCS values in models as compared to observations implies that either the Arctic atmosphere is too cold or has too low an emissivity. Additionally, since larger RLDS differences are found between the models and observations than RLDSCS differences, cloud errors also play a role in this bias.

Additional insight is gleaned about the long-standing bias in RLDS by analyzing the spatial distribution of the bias. Figure 4 shows difference plots of ensemble mean – CERES (a,b) RLDS, (c,d) RLDSCS, and (e,f) LW CRE for January and July, respectively. Although the negative RLDS differences are a domain-wide feature in January, the wintertime underestimate in RLDS is largest in the North Atlantic sector, suggesting that this bias relates to the storm track. The image from July illustrates that while the domain-averaged ensemble mean for July agrees with

CERES, the agreement results from a widespread cancellation between spatial differences. RLDSCS (Fig. 4 c,d) follows the pattern of RLDS in that models simulate too little RLDSCS domain-wide in winter, yet opposing regional biases in summer cancel out and improve the Arctic-average July bias. Since LW CRE is determined by these two fluxes (Eq. 1), the LW CRE biases (Fig. 4 e,f) can be attributed to either RLDS or RLDSCS using the information in Fig. 4. A positive bias in LW CRE (models simulate more longwave cloud forcing than observed) will exist in one of two ways: Either model RLDSCS is too small (such as in January in the Norwegian, Laptev, and Kara Seas or July over the Barents Sea), or model RLDS is too large (such as in July over the Beaufort and Chukchi Seas). A negative LW CRE bias is present when either RLDSCS is too large (January over Greenland) or RLDS is too small (e.g. in January over the central Arctic Ocean). A similar bias in both RLDS and RLDSCS will lead to a LW CRE that has very little difference with observations (such as over Greenland in July; overestimates in both RLDS and RLDSCS approaching 40 Wm^{-2} cancel, leaving a LW CRE biased low by less than 5 Wm^{-2}).

Large intermodel spread is found in the RSDS and RSDSCS seasonal cycle especially during June, July, and August (Fig. 3c). The spatial characteristics of the June, July, and August RSDS and RSDSCS differences closely resemble the annual mean spatial pattern (Fig. 2). In summer, the ensemble mean is biased low over the Arctic Ocean but this difference is offset in the domain average by the large bias over land and the North Atlantic. Larger differences between the models and CERES are found in RSDS as compared to RSDSCS, suggesting that summer time cloud differences have a large influence on the across model RSDS seasonal cycle differences.

A larger spread is found in RSUS and RSUSCS than in the RSDS or RSDSCS indicating that the model disagreement in SW CRE is influenced by differences in surface albedo. Evident from

Fig. 5b, the spread in domain-wide monthly mean surface albedo across models is ~0.2 for all sunlit months and the spatial variability across models is even larger. Figure 6 illustrates the spatial variability of the differences between the ensemble average and CERES observed surface albedo computed using eq. (4). Biases are largest in spring and fall when sea ice begins melting or refreezing. In April and May, an underestimation of CERES SFC-EBAF surface albedo occurs over land approaching -0.4, while the models overestimate albedo in the North Atlantic by up to 0.35. As autumn approaches, a negative bias is most prominent over the Central Arctic with local positive biases present around the coast of Greenland and the Canadian Archipelago. The model surface albedo is influenced by a combination of many factors: namely, 1) differences in model sea ice distribution (including extent and thickness), 2) snow depth on ice, 3) different parameterizations of sea ice albedo [Karlsson and Svensson, 2014], 4) different surface types, such as land with or without snow, or sea ice with or without melt ponds [English et al., 2015], and 5) surface temperature, particularly during melting or freezing, when a small perturbation in surface temperature strongly impacts the physical properties of ice [Koenig et al., 2014].

Figure 7 shows the seasonal cycle of (a) LW CRE, (b) SW CRE, and (c) Net CRE for models and CERES SFC-EBAF. The seasonal cycle of LW CRE matches the shape of the surface downwelling fluxes and is biased low in the winter months. The ensemble average LW CRE in summer agrees well with CERES, however Fig. 7a illustrates a significant spread across the models ranging from 25 and 65 W m^{-2} in July. The annual cycle of cloud fraction (Fig. 5a) explains a portion of the underestimation of LW CRE in winter and the large across-model spread in both winter and summer. The annual mean cloudiness of the models is biased ~12% lower than C3M observations, with wintertime accounting for most of the difference. Models with substantially lower cloud fractions have a dampened cloud greenhouse effect (not shown)

contributing to the underestimation of wintertime LW CRE. Each model in Fig. 5a shows a different shape of the cloud fraction seasonal cycle. Most models indicate minimum cloud fraction during winter and a maximum cloud fraction at end of summer and early fall. Some models, however, simulate more wintertime clouds than summer. Several models capture the observed winter and summer cloud fraction difference but possess a high-amplitude annual cycle with too few wintertime clouds.

Examining the spatial pattern difference in LW CRE is valuable for assessing the spatial contributions to the domain-averaged biases. Figure 4(e-f) shows LW CRE differences between the ensemble and CERES for January and July. The North Atlantic and Greenland regions are the largest contributors to the domain averaged wintertime underestimate in LW CRE. This bias is found to correlate with a model's sea ice cover in this region where too much sea ice coverage results in too little LW CRE (not shown). May (not shown) is the month with the largest underestimate of LW CRE by the ensemble. The feature is very interesting because recent satellite observations suggest an inCREase in spring cloud cover [Wang and Key, 2003, 2005; Schweiger, 2004, Liu et al., 2007]. The inability of models to capture the spring LW CRE and associated cloud properties raises questions about a model's ability to correctly simulate spring cloud trends. Another factor that may influence the seasonality of the LW CRE bias is the timing of the onset of the spring melt season; some models melt earlier while some remain ice-covered, potentially influencing summer cloud conditions [Koenig et al., 2014]. In July (Fig. 4f) when all models melt sea ice and begin producing summertime cloud, they significantly overestimate LW CRE over the Arctic Ocean. The overestimation of summer LW CRE over the Arctic Ocean is common to most models and suggests that the response of cloud to sea ice may be too large.

The annual ensemble average comparison above does not capture the vast intermodel spread in SW CRE during summer. The annual ensemble mean SW CRE is biased low by 3.76 W m^{-2}, however individual models show a summer SW CRE difference of up to 40 Wm^{-2} more negative than CERES while two models show a SW CRE more positive than CERES. A model's simulation of sea ice in summer and the resulting surface albedo change may also explain the large spread in the simulation of summertime SW CRE. Figure 7d shows the SW CRE that models would have if each model possessed the ensemble mean surface albedo. Removing the influence of the model spread in surface albedo brings the models in better agreement with CERES SW CRE; the 90% confidence interval for the difference in means between CERES and models includes the CERES SW CRE for all months in Fig. 7d. The albedo adjustment helps most models produce a SW CRE closer to CERES, particularly for the models greatly underestimating summer SW CRE (e.g BNU-ESM, IPSL-CM5A-MR, BCC-CSM1.1(m)). Three models show no change in SW CRE from the albedo adjustment (e.g. models already possessing a surface albedo close to the ensemble mean, such as CNRM-CM5, ACCESS1.3, INM-CM4), two models show a SW CRE bias that changes sign after the surface albedo adjustment (GFDL-CM3, GISS-E2-R), and some models with larger-than-average albedos that were already simulating a reasonable SW CRE inCREase their biases (e.g. CCSM4, NorESM1-M, MPI-ESM-MR, MPI-ESM-LR). Figure 7e shows the change in SW CRE ($\Delta SW\ CRE$) that results from assuming each model possesses the ensemble mean albedo.

While the ensemble average summer net CRE is consistent with CERES and previous work [e.g., Kay and L'Ecuyer, 2013], large intermodel spread in summer net CRE is found nearly matching the spread in SW CRE during summer. The results (Fig. 7) indicate that 1) several models tend to overestimate the cooling effect and 2) a couple models produce a near zero or

even a positive net CRE in summer. Model differences in surface albedo greatly impact summertime Net CRE; for models with a higher surface albedo, the warming cloud greenhouse effect can more easily exceed the cooling cloud albedo effect. For models with a less reflective surface, the cooling cloud albedo effect is stronger, summarized in Figure 8. A significant correlation ($r = 0.82$) is found between domain-averaged summer net CRE and surface albedo. These findings agree with the results from Karlsson and Svensson, [2013], which found that the summertime disCREpancy in SW CRE is driven by the parameterization of sea ice albedo in CMIP5 models. The models and CERES exhibit the same seasonal cycle amplitude and shape in net CRE, though models underestimate wintertime net CRE by up to 12 W m^{-2}, which follows the biases found in LW CRE due to the absence of SW CRE.

The model-observation differences are analyzed by surface type (not shown). The underestimation of net CRE in the winter occurs primarily over land; wintertime ensemble mean net CRE over land is ~20-25 W m^{-2} lower than CERES while no statistical difference is found over ocean. Despite the wintertime biases stemming from land, the model standard deviation over land during winter is small indicating that all models possess a similar bias. This result indicates that all models may misrepresent the same process in the same manner. The summertime bias in net CRE occurs over ocean with many models simulating a net CRE ~30 W m^{-2} more negative than CERES SFC-EBAF, while in summer net CRE over land in models is consistent with observations.

4. Decomposition of the Cloud Radiative Effect Seasonal Cycle

The seasonal cycle of the CRE is the focus of this section. It is interesting to point out that all models capture the seasonal variation and the summer-winter difference in LW CRE. This is even true of models that produce an inverted annual cycle of total cloud fraction. Thus, further

analysis of the contributions of the annual cycle in LW and SW CRE is necessary to attribute the contributions from cloud fraction and cloud optical property changes. Additionally, Vavrus et al., [2009] found that projected Arctic cloud change is related to characteristics of the simulated cloud seasonal cycle, also serving as motivation to investigate the CRE seasonal cycle.

a. Longwave

Using the independent column approximation, the equation for LW CRE is decomposed into individual components influencing the seasonal cycle: a cloud fraction term and a cloud and clear-sky flux differences term. First, the downwelling all-sky longwave flux is broken up into clear-sky and cloudy-sky components using

$$RLDS = (1 - N) \cdot RLDSCS + N \cdot F \downarrow_{cld,lw} \quad . \qquad (5)$$

$$F \downarrow_{cld,lw} = \frac{(RLDS - (1-N) \cdot RLDSCS)}{N} \qquad (6)$$

In eq. (5), *N* represents total column cloud fraction and $F \downarrow_{cld,lw}$ represents the downwelling LW flux at the surface from an overcast (*N=1*) atmospheric column. Observed *N* is taken from the C3M data set and the observational decomposition is only performed from June 2006-December 2008. Eq. (5) is then substituted into the eq. (1), yielding eq. (7). Lastly, a first-order Taylor Series approximation is performed.

$$LW\ CRE = RLDS - RLDSCS = N \cdot (F \downarrow_{cld,lw} - RLDSCS) \qquad (7)$$

$$\delta LW\ CRE = \delta N \cdot \left(F \downarrow_{cld,lw} - RLDSCS\right) + \delta[\left(F \downarrow_{cld,lw} - RLDSCS\right)] \cdot N \qquad (8)$$

The difference terms denoted by δ in (8) are the difference between a month and the annual mean. The first term in (8) is called the cloud fraction term, δN; it represents the contributions of the seasonal cycle of cloud fraction to the seasonal cycle of LW CRE. The second term is the flux difference term, δF, representing the impact of the difference between cloudy and clear-sky

fluxes on LW CRE, which are primarily due to the emissivity difference resulting from the presence of cloud and can be thought of as cloud optical depth contributions.

Figure 9 shows the annual cycle of (a) δLW CRE, (b) the δN term, and (c) the δF term. The observed δN and δF terms show a different seasonality. The δF term shows larger seasonal-cycle amplitude than δN. The early spring peak in δLW CRE primarily results from contributions of the δF term, whereas the second peak in fall has roughly equal contributions from δN and δF. The ensemble mean δLW CRE matches the C3M observations during wintertime but does not capture δLW CRE in July. Figure 9a also indicates that CMIP5 models simulate too strong of a seasonal cycle in LW CRE. Much like LW CRE, the δLW CRE seasonal cycle also exhibits large intermodel differences throughout the year, though most models share a similar shape.

An anticorrelation is found across models between the δN and δF terms. In other words, models with a larger seasonal cycle in the δN term have a smaller seasonal cycle in δF and vice versa. Additionally, models that simulate the cloud fraction seasonal cycle (Fig. 5a) more similar to C3M observations also exhibit a better representation of the seasonal cycle of the δN term. The relative contributions of the δN and δF terms to the δLW CRE annual cycle seems to be determined by the seasonal cycle simulation of cloud fraction. The better agreement between the models and observation in the δLW CRE seasonal cycle for each model is due to offsetting between the δN and δF terms. The characteristics of the offsetting behavior between the δN and δF terms become more evident when looking at the spatial distribution.

Figure 10 shows the July δLW CRE, δN, and δF terms for three models selected to be illustrative of the intermodel spread in δLW CRE. ACCESS1.0 is an average model representative of the ensemble mean with seasonal cycles of δLW CRE, δN, and δF that match observations reasonably well. It shows a modest amplitude annual cycle of cloud fraction and δN

and δF seasonal cycles that are similar. Each term is of equal importance; this is true in all months for this model. The second model, ACCESS1.3, exhibits an inverted annual cycle of cloud fraction- more cloud in winter than summer. Fig. 9 shows that ACCESS1.3 simulates a δLW CRE that falls within the 90% confidence interval as well, though it exhibits a flat seasonal cycle of δN and a high-amplitude seasonal cycle of δF. In ACCESS1.3 (Fig. 10b), clouds change very little over the year, making the δN term almost negligible. In order to simulate a reasonable seasonal cycle of δLW CRE, the δF term dominates, implying that the cloudy and clear-sky emissivity differences from cloud optical depth changes are simulated to be too strong. Despite these differences, the δLW CRE of ACCESS1.3 very nearly matches that of ACCESS1.0. The third model, CCSM4, has the largest amplitude seasonal cycle of cloud fraction, a large amplitude seasonal cycle of δN, and a flat seasonal cycle of δF. Like the others, it produces a reasonable annual cycle of δLW CRE. Figure 10c shows δN, δF, and δLW CRE for CCSM4; as expected, δN is the dominant term with δF being much smaller. This model relies on changes in cloud amount to drive changes in LW CRE, and less on the seasonality of cloud optical depth. The two terms add together to arrive at a similar δLW CRE spatial distribution in CCSM4 closely matching the other models.

Offsetting biases between the δN and δF terms are also found in other seasons. Figure 11a illustrates the Arctic domain-averaged seasonal cycle of biases (ensemble minus C3M) for δN, δF, and δLW CRE. In winter and early spring, the biases in δN and δF work to offset each other, modulating their impact on δLW CRE. In summertime, positive δN biases over ocean are additive with positive δF biases over land, intensifying the δLW CRE bias. The positive bias in the δN and δF terms in summer is consistent with models producing too strong a cloud response to the exposure of ocean due to melting sea ice. As fall approaches, while biases in δLW CRE

are about the same magnitude as those from spring, the contributions from δN and δF are very different, as are the spatial distribution of these terms. This may be related to the different thermodynamic implications of sea ice melting versus freezing.

b. Shortwave

The SW CRE is also decomposed into individual components using the independent column approximation. Similar to the decomposition of LW CRE, SW CRE can be expressed using cloud fraction, fluxes, and an additional albedo term α. First, RSDS is written using cloud fraction and cloudy and clear-sky fluxes and solved for $F \downarrow_{cld,sw}$.

$$RSDS = N \cdot F \downarrow_{cld,sw} + (1 - N) \cdot RSDSCS \quad (9)$$

$$F \downarrow_{cld,sw} = \frac{RSDS - (1-N) \cdot RSDSCS}{N} \quad (10)$$

Substituting (9) and (10) into the eq. (2) for SW CRE yields an equation expressing SW CRE based on N, α, and shortwave flux differences.

$$SW\ CRE = N \cdot (1 - \alpha) \cdot (F \downarrow_{cld,sw} - RSDSCS) \quad (11)$$

The flux term in the SW CRE decomposition is rewritten as transmission terms using the relations $F \downarrow_{cld} = \tau_{cld} \cdot S$ and $F \downarrow_{clr} = \tau_{clr} \cdot S$, where S is the downwelling shortwave TOA radiation and $\tau_{cld} = \frac{RSDS}{S}$; $\tau_{cld} = \frac{RSDSCS}{S}$. τ_{cld} and τ_{clr} represent the transmissions of solar radiation through cloudy and clear-sky, respectively. This step is done to remove the overwhelming influence of the seasonal cycle of solar insolation from the analysis.

$$SW\ CRE = S \cdot N \cdot (1 - \alpha) \cdot (\tau_{cld} - \tau_{clr}) \quad (10)$$

Applying a first-order Taylor approximation gives the equation for δSW CRE, which contains a cloud fraction term (δN_{SWCRE}), a transmission term ($\delta[(\tau_{cld} - \tau_{clr})]$) (hereafter δτ) , and an albedo term (δα).

$$\delta SW\ CRE = \delta N[S(\tau_{cld} - \tau_{clr})(1 - \alpha)] + \delta(\tau_{cld} - \tau_{clr})[SN(1 - \alpha)] - \delta\alpha[SN(\tau_{cld} - \tau_{clr})]$$

(11)

Recall that a change in α with no change in cloud yields an apparent change in the cloud because of a change in SW CRE. This decomposition allows the effects of α changes on SW CRE to be separated from the effects of cloud changes. δSW CRE as well as the three terms are shown in Fig. 11b, calculated for sunlit months (monthly mean solar insolation > 50 W m^{-2}).

The observed δSW CRE peaks in April and is at its lowest value in August, following the trend in δτ. Throughout the sunlit months, δτ and δα show strong seasonal changes while the observed δN_{SWCRE} is a smaller influence on δSW CRE. The ensemble mean is consistent with observed δSW CRE most of the year but with slightly higher values in spring and slightly lower values in fall. Looking at the three terms (Fig. 12b-d), the spring overestimation is primarily due to the δτ term. The δN term shows a low bias in the summer, though opposite sign biases in δτ and δα cancel out this bias in δSW CRE. The fall underestimation is mostly due to δα with a small contribution from δτ. The same offsetting behavior between terms found in the δLW CRE seasonal cycles is also found in δSW CRE; models with high-amplitude annual cycles of cloud fraction have larger annual cycles of δN_{SWCRE} and flatter annual cycles of δτ and vice versa for the models with low-amplitude annual cycles of cloud fraction. Unlike δN_{SWCRE} and δτ, δα is more consistent between models, though large biases with observations are still present.

Figure 12b shows ensemble mean – C3M δSW CRE biases throughout the annual cycle and the relative contributions from biases in δN_{SWCRE}, δτ, and δα. Most of the year, δτ and δα present a same-sign bias that counteracts the bias of δN_{SWCRE}. In terms of magnitude, the δN_{SWCRE} bias is the largest-magnitude contribution to δSW CRE biases during April, May, July and August. The δN_{SWCRE} bias is positive in spring when models show too much sea ice and too few clouds and is negative in summer when models simulate too much cloud cover over ocean. The effects of

biases in $\delta\alpha$ are strongest in fall, particularly October when sea ice is refreezing, while the bias in $\delta\tau$ is most impactful in March.

5. Discussion

The broader question this study addresses is determining the impact of the model errors in surface fluxes on model simulation of present day Arctic surface temperature. Intuitively, changes in the surface radiative fluxes should lead to changes in surface temperature, but the relation is more complex and depends on surface type, geographic region, season, and dynamical effects, as well as other processes that may be buffering the impact of surface flux biases.

If a surface temperature bias is solely cloud related, we expect a positive correlation between errors in CRE and surface temperature biases. No correlation may indicate that the temperature bias is related to errors in the clear-sky fluxes or compensated for by other processes such as surface turbulent fluxes and the influence of dynamic energy transport by the atmosphere or the ocean. Correlations are computed across the climate models, meaning that a positive correlation between net CRE and surface temperature bias indicates that models with a larger net CRE have a larger average surface temperature. In January, errors in surface temperature over the North Atlantic and the Bering Strait are correlated with errors in net CRE (Fig. 13a). Outside of the East Greenland/Barents Sea and the Bering Strait, winter cloudiness is low and net CRE errors are not strongly correlated with T_s biases. Errors in RLDS and RLDSCS are strongly correlated (generally > 0.5) to T_s biases across the Arctic domain. Model errors in RLDSCS exhibit the largest correlations with T_s errors from December-February. Larger cloud cover in March and November contributes to RLDS errors showing stronger correlations with T_s bias.

As spring approaches and solar radiation reaching the surface increases, model errors in April CRE over the North Atlantic remain strongly correlated to T_s biases though the correlation is

negative, opposite to the expected relationship (Fig. 13b). This indicates that as net CRE becomes more positive, surface temperature becomes colder and another radiative or dynamical process is affecting the relationship. The seasonal shift in the correlation between net CRE and T_s implies a correlation between clouds and advective forcing. While errors in RLDS in April are the best correlated parameter to T_s biases over the central Arctic Ocean and Northern Canada, errors in RSUSCS and RSUS negatively correlate with the T_s biases in April. The negative correlation between RSUS, RSUSCS and T_s bias last throughout summer until October when sea ice begins to refreeze. The negative correlation between RSUS, RSUSCS, and T_s indicates that models with more surface reflection are colder. While this is the physical relationship one would expect based upon the surface energy budget equation, not all radiative flux terms exhibit the expected direct relationship—e.g., RSDS in July (Fig. 13c). The correlation between RSUS, RSUSCS, and T_s are the some of the largest values over the Arctic Ocean. These large correlations are likely because the errors represent differences in α, and the surface albedo feedback is one of the dominant feedbacks controlling Arctic T_s. The correlation between RSUSCS biases and T_s biases is strongest in the regions where surface albedo strongly varies: over cloudy conditions in the North Atlantic, over melting snow and melt ponds, and over sea ice of varying thickness and age. The maximum snow depth, typically observed in May [Serreze and Barry, 2014] coincides with the high correlation between errors in RSUSCS and T_s over Siberian land and Northern Alaska, where models fail to simulate the high albedo (> 0.7) of fresh snow. The strong correlation also occurs when the minimum annual snow depth is achieved in August; this time over the central Arctic sea ice pack, representing model difficulties in simulating sea ice albedo when the ice packs develop dark melt ponds. During the summer, the main shortcoming

of the models is determining surface albedo when one region of sea ice could contain fresh snowfall (α≈0.7-0.9), melting snow (α≈0.5-0.6), and melt ponds (α≈0.15-0.4).

The correlation between RSDS biases and T_s biases are only directly correlated during summer over land (Fig. 13c). Errors in July RSDS over land are strongly correlated (>0.8) to T_s biases. The strong correlation corresponds with the large model standard deviation in summertime RSDS over land. Errors in shortwave fluxes are the most important to determining summertime cloud forcing. The correlations between July net CRE and T_s biases (Fig. 13c) strongly resemble the spatial patterns for RSDS and T_s. Physically, T_s biases are expected to exhibit the direct relationships as implied by the surface energy budget equation over land owing to its lower heat capacity than ocean and lack of a dynamic heat transport mechanism. After the refreezing of sea ice in October (Fig. 13d), Arctic surface temperature biases return to being driven by RLDS (particularly over land) and RLDSCS (over ocean). Errors in October Net CRE show the strongest positive correlations with T_s bias over the Canadian Archipelago and Northern Russia.

The link between across model spread in CRE and winter surface temperature has been studied previously [Karlsson and Svennson, 2011]. The conclusions from that study using the CMIP3 dataset are in agreement with the results presented here; no relationship is found between a model's domain-averaged net CRE and domain-averaged surface temperature. This is counterintuitive given the importance of clouds to the surface energy balance. The analysis presented above provides additional information about the relationship between net CRE and surface temperature by exploring the relationship spatially. We find that the expected positive correlation does exist regionally on a monthly time scale with each month possessing very different spatial distributions. Jul-Aug and Oct-Dec net CRE and T_s are positively correlated

over land while Nov-Feb shows high correlation over the North Atlantic. When studying this relationship using domain-averaged values, the anti-correlation frequently found over the central Arctic Ocean leads to no obvious dependency of surface temperature on net CRE. Strong anti-correlation between these two parameters is more prevalent during spring and summer, and is indicative that another non-cloud related process is controlling surface temperature. The strong regional variability in the correlations between net CRE and T_s, both positive and negative, explain the lack of domain wide correlation found by Karlsson and Svennson, [2011]. Model surface fluxes can explain some of the model spread in surface temperature (e.g. spring/summer RSUS and RSUSCS; winter RLDS and RLDSCS), but regions in which no significant correlation exists between either surface fluxes or cloud forcing and temperature suggest that large-scale dynamics and transport may control model surface temperature for these areas.

6. Conclusions

This paper provides an evaluation of Arctic surface radiation budget in the CMIP5 historical experiment against observations from CERES SFC-EBAF and C3M data products. The main conclusions from the study are summarized below.

- The CMIP5 ensemble mean annually averaged longwave and shortwave all-sky downwelling fluxes (RLDS, RSDS), shortwave all-sky upwelling flux (RSUS), and longwave clear-sky downwelling flux (RLDSCS) are all significantly lower than CERES SFC-EBAF observations. Ensemble mean shortwave clear-sky fluxes however (RSDSCS, RSUSCS) are larger the than observations. Smaller-than-observed downwelling longwave fluxes are most prominent over the North Atlantic. The largest disagreement between the models' radiative flux terms is found in the North Atlantic due to varying representations of cloudiness.

- Improving model simulation of surface albedo, particularly over sea ice, snow, and ice-water boundaries is needed to reduce biases in RSUSCS and SW CRE. While the seasonal cycle of albedo varies between models by ~0.2 during all months for domain-wide averages, regional biases in albedo exceed this value particularly during melting (spring) and freezing (autumn). This result suggests that the rate of sea ice melting during the transition season plays a significant role in surface radiative flux errors.
- Seasonal cycles of LW, SW, and net CRE in CMIP5 are similar to those from CMIP3 [Karlsson and Svensson, 2011]; models underestimate net CRE in the winter due to a domain-wide pattern of smaller than observed LW CRE. Summer net CRE shows a shape and model spread resembling that of summer SW CRE.
- The simulation of wintertime cloud fraction by the models is a large source of error in computing cloud radiative effect. Total column cloud fraction for the models ranges from 30%-95% in winter. The annual average ensemble mean winter cloud fraction is found to be 12% less than combined CALIPSO-CloudSAT derived cloud fraction. Additionally, the various model representations of the cloud fraction seasonal cycle divides the models into two groups that behave radiatively different: models with inverted cycles of cloud fraction with more winter clouds than summer, and models simulating a curve closer to observations with more summer cloud cover than in winter.
- An independent column approximation can be used to decompose LW and SW CRE and attribute variations in the seasonal cycle to individual components. The models that simulate the seasonal cycle of cloud fraction closer to observations are found to have the largest contribution to the seasonal cycle of cloud radiative effect from cloud fraction, while models showing an inverted or flat cloud fraction seasonal cycle exhibit the largest

contributions from changes in the difference between cloudy and clear-sky fluxes mostly due to changes in cloud optical depth. Changes in SW CRE for most models are strongly impacted by changes in surface albedo, while the effects of cloud fraction and the differences in clear- and cloudy-sky transmittance are similarly important to individual models based upon their annual cycles of cloud fraction; those with inverted cycles are more driven by differences in clear and cloudy sky transmittance and vice versa.

- Model biases in surface fluxes are correlated with model biases in surface temperature for certain regions and different seasons. Annually, the relationship between these two quantities is weaker as surface flux biases have a different relationship with temperature biases in each month. The strong correlations presented above are not observed when averaging over the entire Arctic cap. The connection between surface flux errors and temperature biases is tied to surface type: in the same month, errors in a single flux are often positively correlated to T_s biases over land but negatively correlated over ocean and vice versa. This land/ocean contrast is a visible feature when looking at model biases in both surface fluxes and cloud radiative effect.
- The largest errors and across-model spread in the surface radiation budget is found in the North Atlantic. Therefore, future studies should focus on understanding and correcting these errors because it will lead to significant improvements in the Arctic surface radiation budget. Further, since the errors in this region are likely related to the storm track it is likely that this will influence and improve the simulation of the mid-latitude climate and lead to a more realistic coupling between the Arctic and mid-latitudes as the response of the storm track is one of the pathways through which the rapid changes in the Arctic will influence the mid-latitudes.

Acknowledgements. This work is funded by the NASA Interdisciplinary Studies Program grant NNH12ZDA001N-IDS. The processing of the C3M data used in this analysis was funded under the NASA Energy and Water Cycle Studies program. The C3M and CERES data is available from the Langley Atmospheric Science Data Center (http://eosweb.larc.nasa.gov).

7. References

Chylek, P., C. K. Folland, G. Lesins, M. K. Dubey, and M. Wang (2009), Arctic air temperature change amplification and the Atlantic Multidecadal Oscillation, *Geophys. Res. Lett. Geophysical Research Letters*, *36*(14), doi:10.1029/2009gl038777.

English, J. M., A. Gettelman, and G. R. Henderson (2015), Arctic Radiative Fluxes: Present-Day Biases and Future Projections in CMIP5 Models, *Journal of Climate J. Climate*, *28*(15), 6019–6038, doi:10.1175/jcli-d-14-00801.1.

Graversen, R. G., and M. Wang (2009), Polar amplification in a coupled climate model with locked albedo, *Clim Dyn Climate Dynamics*, *33*(5), 629–643, doi:10.1007/s00382-009-0535-6.

Hall, A. (2004), The Role of Surface Albedo Feedback in Climate, *Journal of Climate J. Climate*, *17*(7), 1550–1568, doi:10.1175/1520-0442(2004)017<1550:trosaf>2.0.co;2.

Hall, A., and X. Qu (2006), Using the current seasonal cycle to constrain snow albedo feedback in future climate change, *Geophys. Res. Lett. Geophysical Research Letters*, *33*(3), doi:10.1029/2005gl025127.

Holland, M. M., and C. M. Bitz (2003), Polar amplification of climate change in coupled models, *Climate Dynamics*, *21*(3-4), 221–232, doi:10.1007/s00382-003-0332-6.

Karlsson, J., and G. Svensson (2013), Consequences of poor representation of Arctic sea-ice albedo and cloud-radiation interactions in the CMIP5 model ensemble, *Geophys. Res. Lett. Geophysical Research Letters*, *40*(16), 4374–4379, doi:10.1002/grl.50768.

Karlsson, J., and G. Svensson (2010), The simulation of Arctic clouds and their influence on the winter surface temperature in present-day climate in the CMIP3 multi-model dataset, *Clim Dyn Climate Dynamics*, *36*(3-4), 623–635, doi:10.1007/s00382-010-0758-6.

Kato, S., et al. (2010), Relationships among cloud occurrence frequency, overlap, and effective thickness derived from CALIPSO and CloudSat merged cloud vertical profiles, *J. Geophys. Res., 115, D00H28*, doi: 10.1029/2008JD012277.

Kato, S. et al. (2011), Improvements of top-of-atmosphere and surface irradiance computations with CALIPSO-, CloudSat-, and MODIS-derived cloud and aerosol properties, *J. Geophys. Res. Journal of Geophysical Research*, *116*(D19), doi:10.1029/2011jd016050.

Kato, S. et al. (2013), Surface irradiances consistent with CERES-derived top-of-atmosphere shortwave and longwave irradiances, *J. Climate*, *26*, 2719-2740, doi: 10.1175/JCLI-D-12-00436.1.

Kay, J. E., and A. Gettelman (2009), Cloud influence on and response to seasonal Arctic sea ice loss, *J. Geophys. Res. Journal of Geophysical Research*, *114*(D18), doi:10.1029/2009jd011773.

Kay, J. E., and T. L'ecuyer (2013), Observational constraints on Arctic Ocean clouds and radiative fluxes during the early 21st century, *Journal of Geophysical Research: Atmospheres J. Geophys. Res. Atmos.*, *118*(13), 7219–7236, doi:10.1002/jgrd.50489.

Koenigk, T., A. Devasthale, and K.-G. Karlsson (2014), Summer Arctic sea ice albedo in CMIP5 models, *Atmospheric Chemistry and Physics Atmos. Chem. Phys.*, *14*(4), 1987–1998, doi:10.5194/acp-14-1987-2014.

Manabe, S., and R. T. Wetherald (1975), The Effects of Doubling the CO 2 Concentration on the climate of a General Circulation Model, *J. Atmos. Sci. Journal of the Atmospheric Sciences*, *32*(1), 3–15, doi:10.1175/1520-0469(1975)032<0003:teodtc>2.0.co;2.

Sejas, S, M. Cai, A. Hu, J. Meehl, W. Washington, and P. C. Taylor (2014), On the seasonality of polar warming amplification, *J. Climate*, *27*, 5653-5669.

Serreze, M. C., and R. G. Barry (2005), *The Arctic climate system*, Cambridge University Press, Cambridge.

Sorteberg, A., V. Kattsov, J. E. Walsh, and T. Pavlova (2007), The Arctic surface energy budget as simulated with the IPCC AR4 AOGCMS, *Clim. Dyn.* 29, 131-156, doi:10.1007/s00382-006-0222-9.

Stephens, G. L. et al. (2008), CloudSat mission: Performance and early science after the first year of operation, *J. Geophys. Res. Journal of Geophysical Research*, *113*, doi:10.1029/2008jd009982.

Taylor, K. E., R. J. Stouffer, and G. A. Meehl (2012), An Overview of CMIP5 and the Experiment Design, *Bull. Amer. Meteor. Soc. Bulletin of the American Meteorological Society*, *93*(4), 485–498, doi:10.1175/bams-d-11-00094.1.

Taylor, P. C., R. G. Ellingson, and M. Cai (2011), Geographic distribution of climate feedbacks in the NCAR CCSM3.0, *J. Climate*, *24*, 2737-2753. Doi: http://dx.doi.org/10.1175/2010JCLI3788.1.

Taylor, P. C., R. G. Ellingson, and M. Cai (2011), Seasonal distribution of climate feedbacks in the NCAR CCSM3.0, *J. Climate*, *24*, 3433-3444. Doi: http://dx.doi.org/10.1175/2011JCLI3862.1.

Taylor, P. C., M. Cai, A. Hu, J. Meehl, W. Washington, G. J. Zhang (2013), A Decomposition of Feedback Contributions to Polar Warming Amplification, *J. Climate*, *26*, 7023–7043. doi: http://dx.doi.org/10.1175/JCLI-D-12-00696.1.

Taylor, P. C., S. Kato, K.-M. Xu, and M. Cai (2015), Covariance between Arctic sea ice and clouds within atmospheric state regimes at the satellite footprint level, *J. Geophys. Res. Atmos.*, *120*, 12656-12678. doi:10.1002/2015JD023520.

Vavrus, S. J., U. S. Bhatt, and V. A. Alexeev (2011), Factors Influencing Simulated Changes in Future Arctic Cloudiness, *Journal of Climate J. Climate*, *24*(18), 4817–4830, doi:10.1175/2011jcli4029.1.

Vavrus, S., D. Waliser, A. Schweiger, and J. Francis (2008), Simulations of 20th and 21st century Arctic cloud amount in the global climate models assessed in the IPCC AR4, *Clim Dyn Climate Dynamics*, *33*(7-8), 1099–1115, doi:10.1007/s00382-008-0475-6.

Wielicki, B. A., B. R. Barkstrom, E. F. Harrison, R. B. Lee, G. L. Smith, and J. E. Cooper (1996), Clouds and the Earth's Radiant Energy System (CERES): An Earth Observing System Experiment, *Bull. Amer. Meteor. Soc. Bulletin of the American Meteorological Society*, *77*(5), 853–868, doi:10.1175/1520-0477(1996)077<0853:catere>2.0.co;2.

Winker, D. M. et al. (2010), The CALIPSO Mission: A Global 3D View of Aerosols and Clouds, *Bull. Amer. Meteor. Soc. Bulletin of the American Meteorological Society*, *91*(9), 1211–1229, doi:10.1175/2010bams3009.1.

Table Captions

Table 1. List of CMIP5 models used and country of origin.

Table 2. Summary of average surface radiative fluxes for each model and CERES Surface EBAF observations.

Table 3. Summary of average longwave, shortwave, and net cloud radiative effect for each model and CERES Surface EBAF observations.

Figure Captions

Figure 1. Annual mean spatial distribution of the surface radiative fluxes for CERES SFC-EBAF observations(left column), Ensemble Mean – CERES SFC-EBAF(middle column), and Ensemble standard deviation(right column): (a-c) RLDS, (d-f) RLDSCS, (g-i) RSDS, (j-l) RSDSCS, (m-o) RSUS, and (p-r) RSUSCS.

Figure 2. Annual mean spatial distribution of the surface cloud radiative effect terms for CERES SFC-EBAF observations (left column), Ensemble Mean (middle column), and Ensemble standard deviation (right column): (a-c) LW CRE, (d-f) SW CRE, and (g-i) Net CRE.

Figure 3. Arctic domain average—latitude > 66°N—seasonal cycle for (a) RLDS, (b) RLDSCS, (c) RSDS, (d) RSDSCS, (e) RSUS), and (f) RSUSCS.

Figure. 4. Difference plots of the Ensemble mean minus CERES RLDS for (a) January and (b) July; Ensemble mean minus CERES RLDSCS for (c) January and (d) July, and Ensemble mean minus CERES LW CRE for (e) January and (f) July.

Figure 5. Arctic domain average—latitude > 66°N—seasonal cycle of (a) cloud fraction and (b) surface albedo from observations and CMIP5 models. Cloud fraction annual cycle is from the C3M data set using active remote sensing. Albedo annual cycle is from CERES SFC-EBAF.

Figure 6. Annual cycle (April through September) of surface albedo bias in the CMIP5 Ensemble average minus CERES SFC-EBAF.

Figure 7. Arctic domain average—latitude > 66°N—seasonal cycle of (a) LW CRE, (b) SW CRE, (c) net CRE, (d) Albedo-adjusted SW CRE, and (e) Albedo contributions to SW CRE from observations and CMIP5 models.

Figure 8. June-July-August domain-averaged Net CRE vs JJA surface albedo.

Figure. 9. Arctic domain-average seasonal cycle of the LW CRE decomposition terms (a) δLW CRE, (b) δN, and (c) δF. Units are W m^{-2}.

Figure 10. Spatial distributions of the LW CRE decomposition terms for (a) ACCESS1.0, (b) ACCESS1.3, and (c) CCSM4.

Figure 11. Arctic domain-average biases (Ensemble mean minus observations) for (a) LW CRE decomposition terms, and (b) SW CRE decomposition terms.

Figure 12. Arctic domain-average seasonal cycle of the SW CRE decomposition terms (a) δSW CRE, (b) δN, (c) $\delta\tau$ and d) -$\delta\alpha$. Units are W m^{-2}.

Figure 13. Spatial distributions of correlation coefficients between surface radiative flux biases and surface temperature biases for (a) January, (b) April, (c) July, and (d) October.

Table 1. List of CMIP5 models used and country of origin.

Model	Country
ACCESS1.0	Australia
ACCESS1.3	Australia
BCC-CSM1.1	China
BCC-CSM1.1(m)	China
BNU-ESM	China
CCSM4	USA
CNRM-CM5	France
CSIRO-Mk3.6.0	Australia
GFDL-CM3	USA
GISS-E2-R	USA
INM-CM4	Russia
IPSL-CM5A-MR	France
MIROC5	Japan
MPI-ESM-MR	Germany
MPI-ESM-LR	Germany
MRI-CGCM3	Japan
NorESM1-M	Norway

Table 2. Summary of average surface radiative fluxes for each model and CERES Surface EBAF observations.

	RLDS	RLDSCS	RSDS	RSDSCS	RSUS	RSUSCS
ACCESS1.0	218.61	184.57	97.43	132.55	51.85	67.66
ACCESS1.3	222.46	186.74	96.54	132.79	55.56	72.40
BCC-CSM1.1	225.72	182.63	84.51	130.46	42.67	60.75
BCC-CSM1.1(m)	226.57	188.96	86.12	130.48	42.45	58.45
BNU-ESM	220.82	185.81	82.89	129.84	48.11	68.91
CCSM4	216.31	185.93	93.27	131.18	59.42	79.55
CNRM-CM5	209.12	182.78	104.5	124.52	61.58	70.32
CSIRO-Mk3.6.0	212.93	170.45	95.18	136.00	59.86	82.54
GFDL-CM3	228.93	191.90	90.08	133.61	53.86	75.65
GISS-E2-R	208.46	183.34	105.46	131.73	45.92	56.42
INMCM4	231.46	188.84	89.98	124.33	48.81	65.13
IPSL-CM5A-LR	211.41	177.44	88.58	133.99	45.96	67.61
IPSL-CM5A-MR	222.37	184.36	85.03	133.20	40.57	61.28
MIROC5	216.75	182.34	104.19	134.77	68.93	85.87
MPI-ESM-LR	236.22	189.00	91.15	140.49	57.00	83.00
MPI-ESM-MR	234.30	189.78	93.74	139.39	57.99	81.45
MRI-CGCM3	208.95	182.28	99.02	135.65	57.04	73.49
NorESM1-M	217.51	183.08	86.74	130.00	55.35	78.42
Ensemble Mean	220.5 ± 8.7	184.5 ± 4.9	93.0 ± 7.1	132.5 ± 4.2	52.9 ± 7.7	71.6 ± 9.0
CERES-EBAF	230.45	190.69	95.92	127.64	55.77	68.85

Table 3. Summary of average longwave, shortwave, and net cloud radiative effect for each model and CERES Surface EBAF observations.

	LW CRE	SW CRE	Net CRE
ACCESS1.0	34.04	-18.40	15.64
ACCESS1.3	35.72	-17.62	18.10
BCC-CSM1.1	43.09	-25.81	17.28
BCC-CSM1.1(m)	37.61	-26.11	11.50
BNU-ESM	35.01	-23.90	11.11
CCSM4	30.39	-16.40	13.99
CNRM-CM5	26.34	-9.55	16.79
CSIRO-Mk3.6.0	42.48	-16.68	25.80
GFDL-CM3	40.05	-19.08	20.96
GISS-E2-R	24.61	-16.56	8.06
INMCM4	42.62	-17.19	25.43
IPSL-CM5A-LR	33.97	-23.45	10.52
IPSL-CM5A-MR	38.01	-26.90	11.11
MIROC5	34.41	-12.24	22.17
MPI-ESM-LR	47.23	-21.22	26.00
MPI-ESM-MR	44.53	-20.03	24.49
MRI-CGCM3	26.67	-18.56	8.11
NorESM1-M	34.44	-18.04	16.39
Ensemble Mean	36.18 ± 6.48	-19.32 ± 4.65	16.86 ± 6.14
CERES-EBAF	39.78	-15.56	24.22

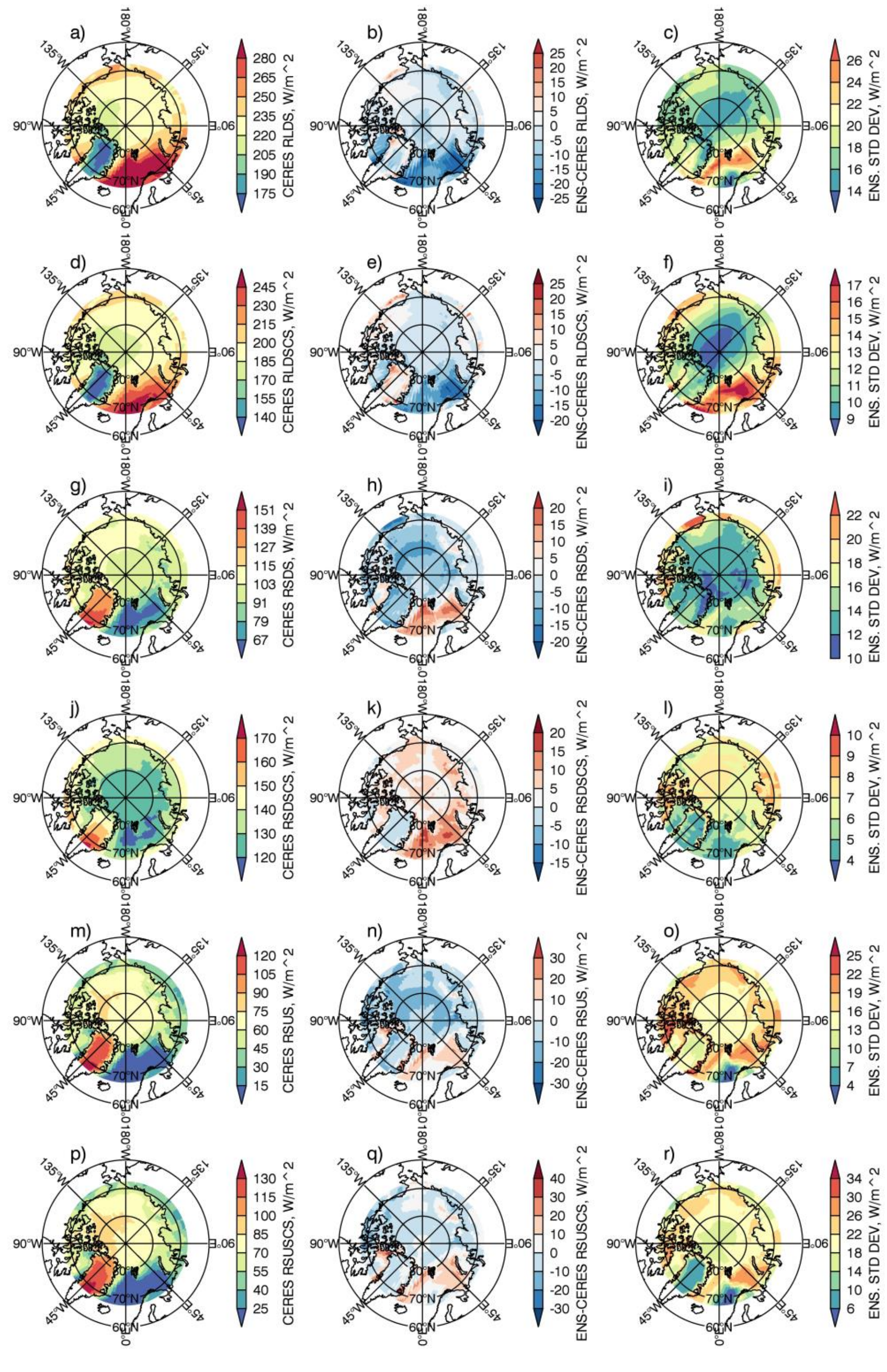

Figure 1. Annual mean spatial distribution of the surface radiative fluxes for CERES SFC-EBAF observations (left column), Ensemble Mean – CERES SFC-EBAF (middle column), and Ensemble standard deviation (right column): (a-c) RLDS, (d-f) RLDSCS, (g-i) RSDS, (j-l) RSDSCS, (m-o) RSUS, and (p-r) RSUSCS.

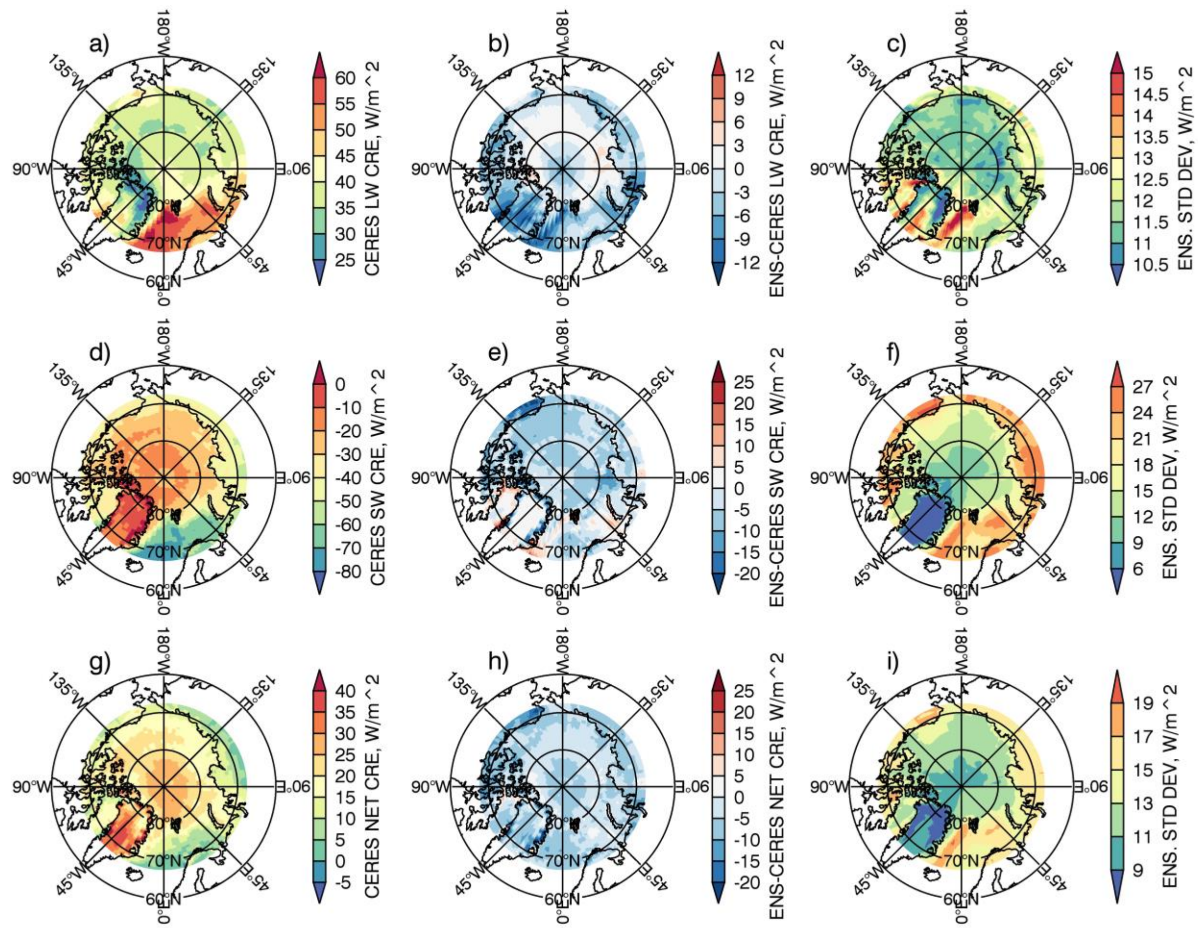

Figure 2. Annual mean spatial distribution of the surface cloud radiative effect terms for CERES SFC-EBAF observations (left column), Ensemble Mean (middle column), and Ensemble standard deviation (right column): (a-c) LW CRE, (d-f) SW CRE, and (g-i) Net CRE.

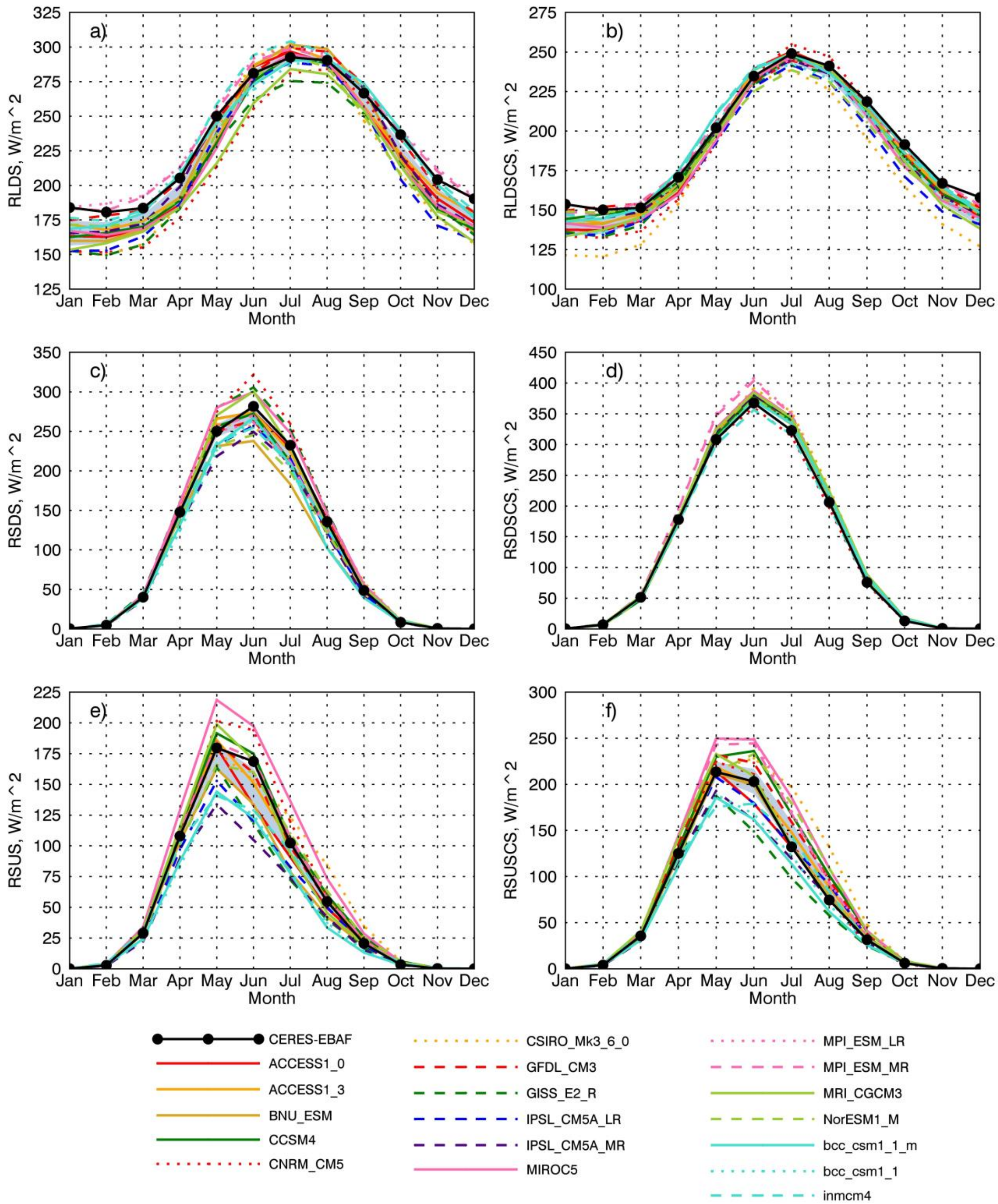

Figure 3. Arctic domain average—latitude > 66°N—seasonal cycle for (a) RLDS, (b) RLDSCS, (c) RSDS, (d) RSDSCS, (e) RSUS), and (f) RSUSCS.

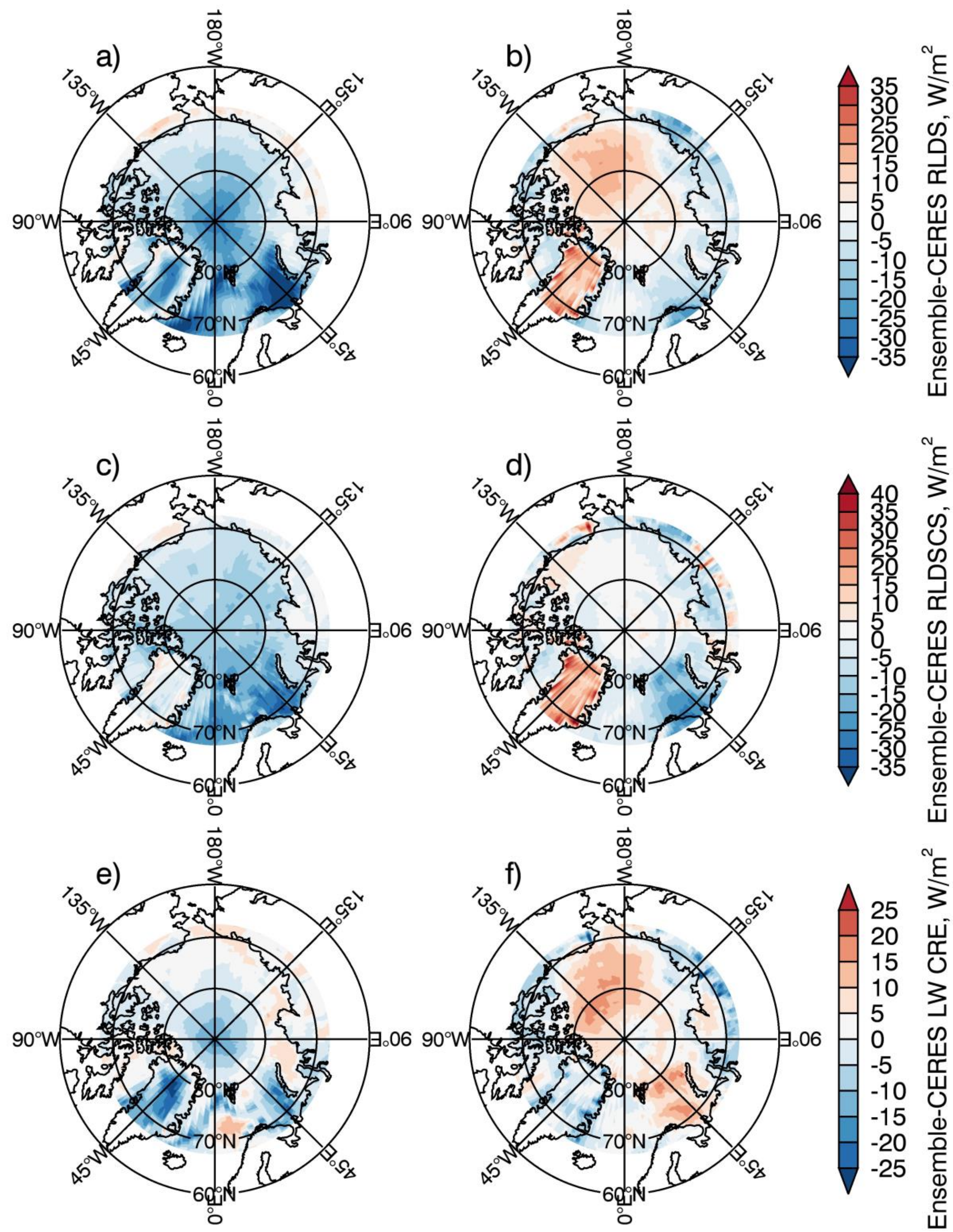

Figure. 4. Difference plots of the Ensemble mean minus CERES RLDS for (a) January and (b) July; Ensemble mean minus CERES RLDSCS for (c) January and (d) July, and Ensemble mean minus CERES LW CRE for (e) January and (f) July.

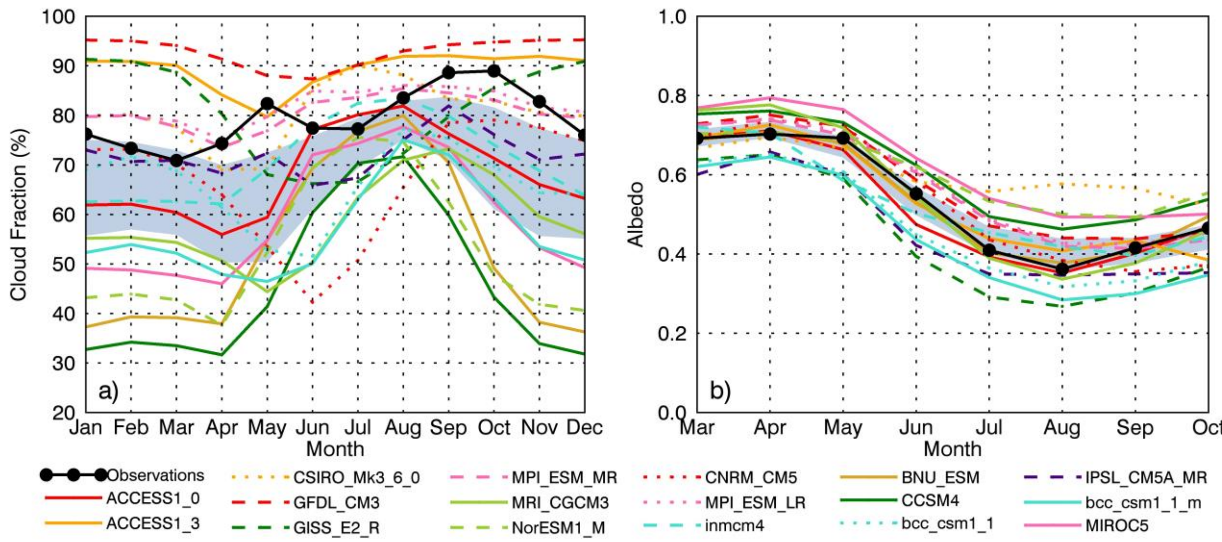

Figure 5. Arctic domain average—latitude > 66°N—seasonal cycle of (a) cloud fraction and (b) surface albedo from observations and CMIP5 models. Cloud fraction annual cycle is from the C3M data set using active remote sensing. Albedo annual cycle is from CERES SFC-EBAF.

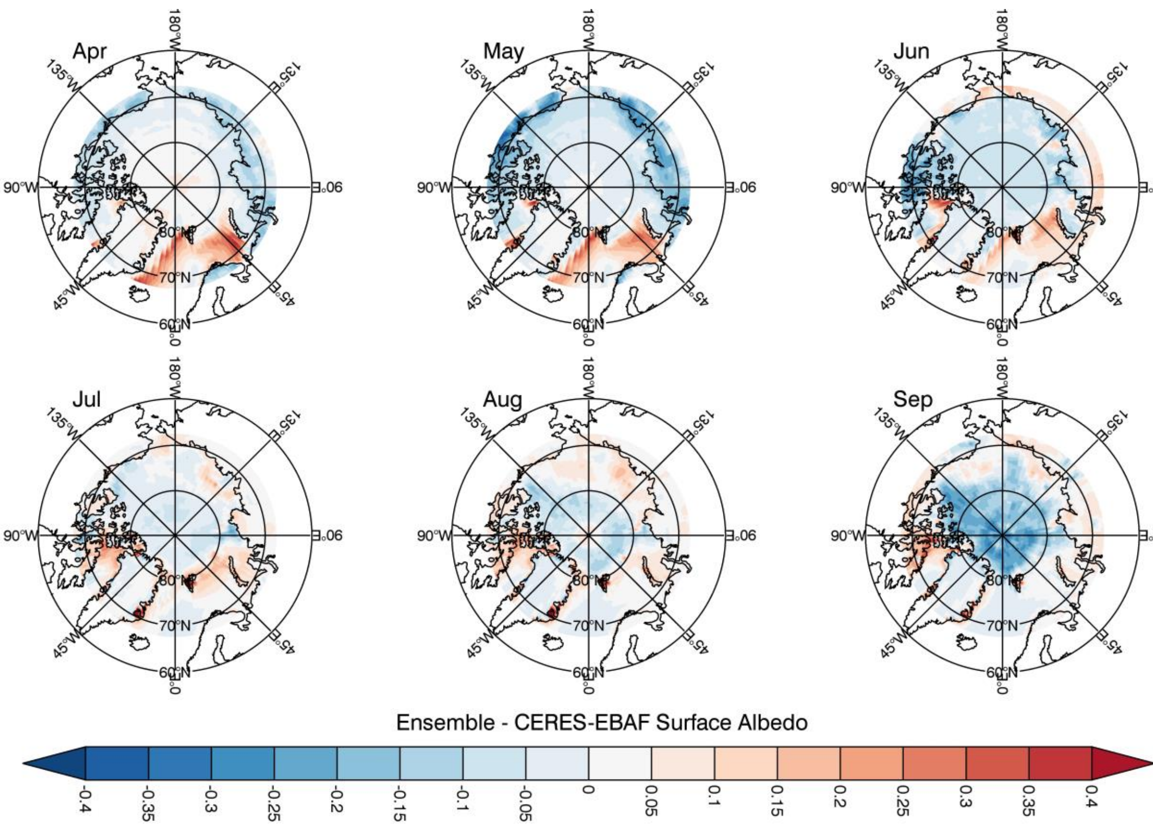

Figure 6. Annual cycle (April through September) of surface albedo bias in the CMIP5 Ensemble average minus CERES SFC-EBAF.

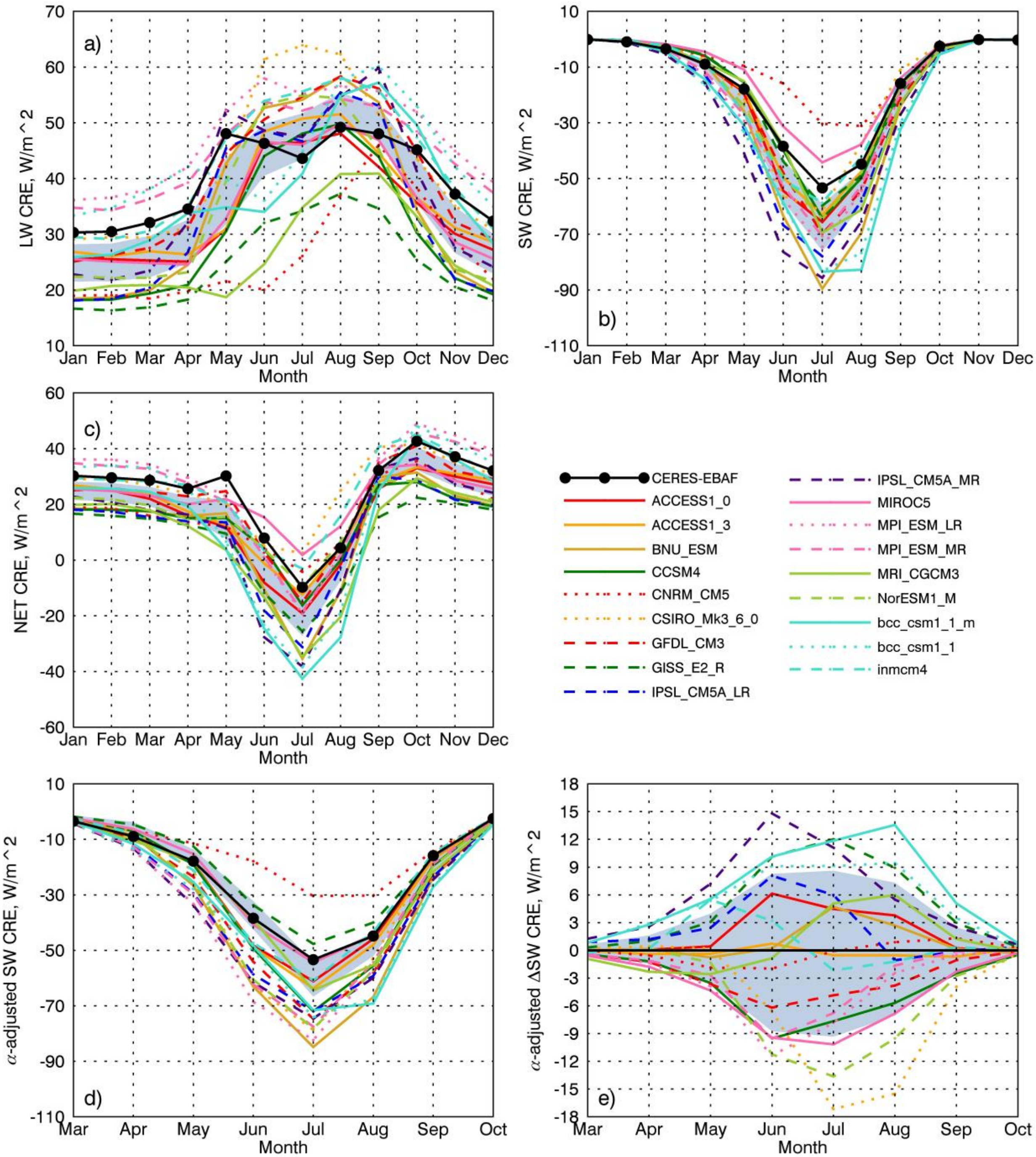

Figure 7. Arctic domain average—latitude > 66°N—seasonal cycle of (a) LW CRE, (b) SW CRE, (c) net CRE, (d) Albedo-adjusted SW CRE, and (e) Albedo contributions to SW CRE from observations and CMIP5 models.

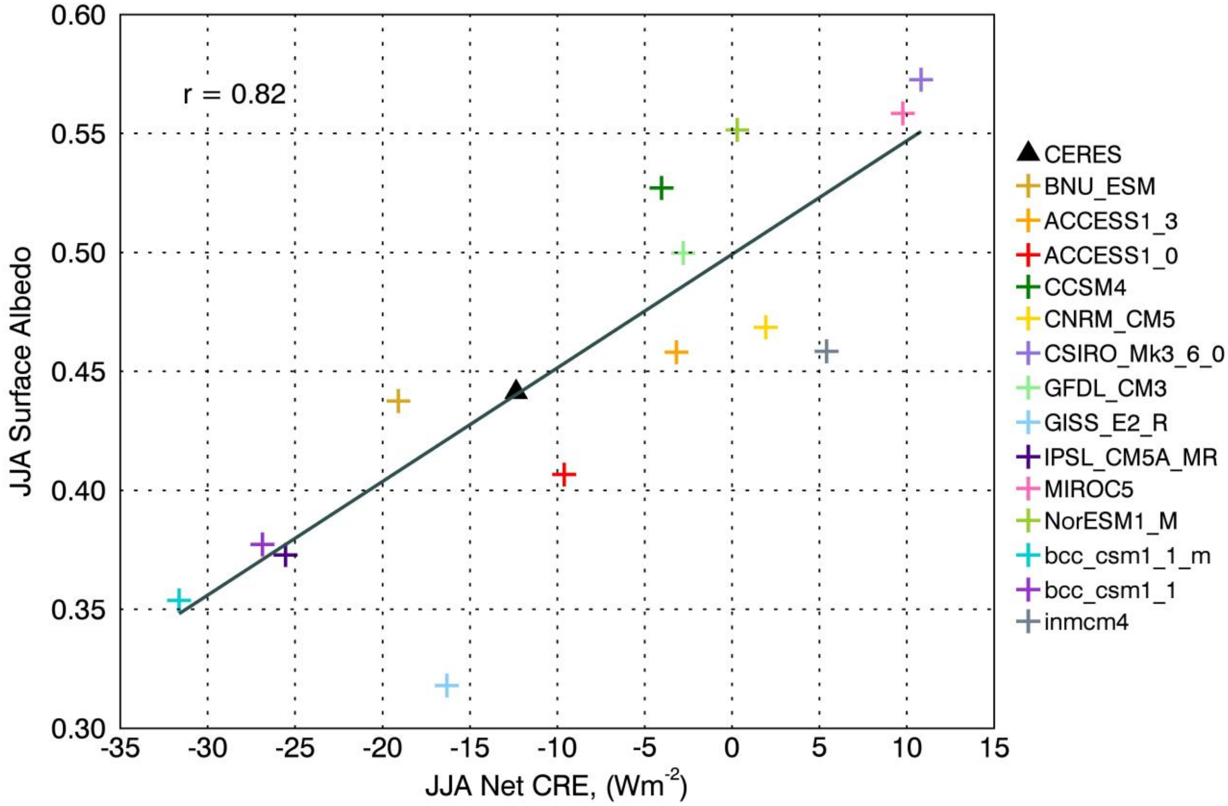

Figure 8. June-July-August domain-averaged Net CRE vs. JJA surface albedo.

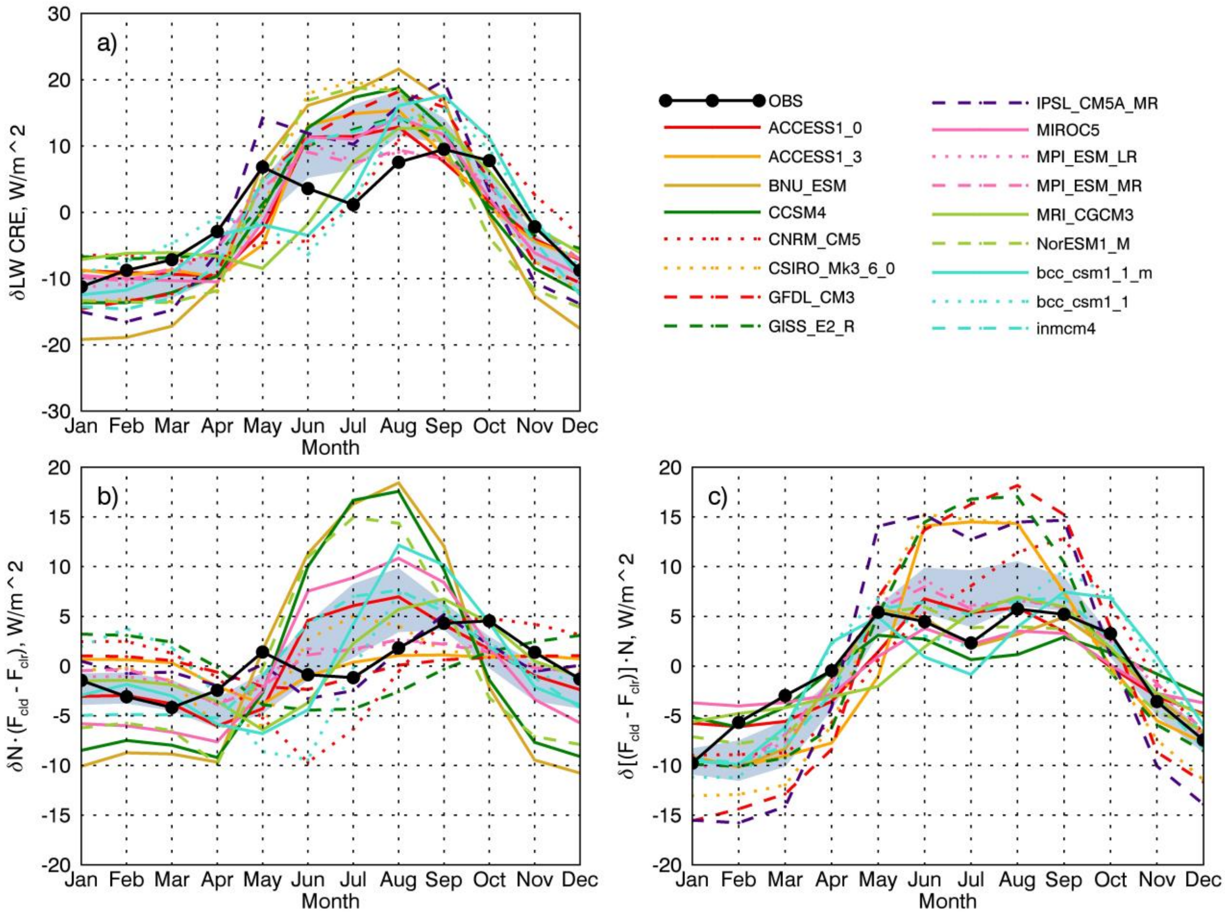

Figure. 9. Arctic domain-average seasonal cycle of the LW CRE decomposition terms (a) δLW CRE, (b) δN, and (c) δF. Units are W m^{-2}.

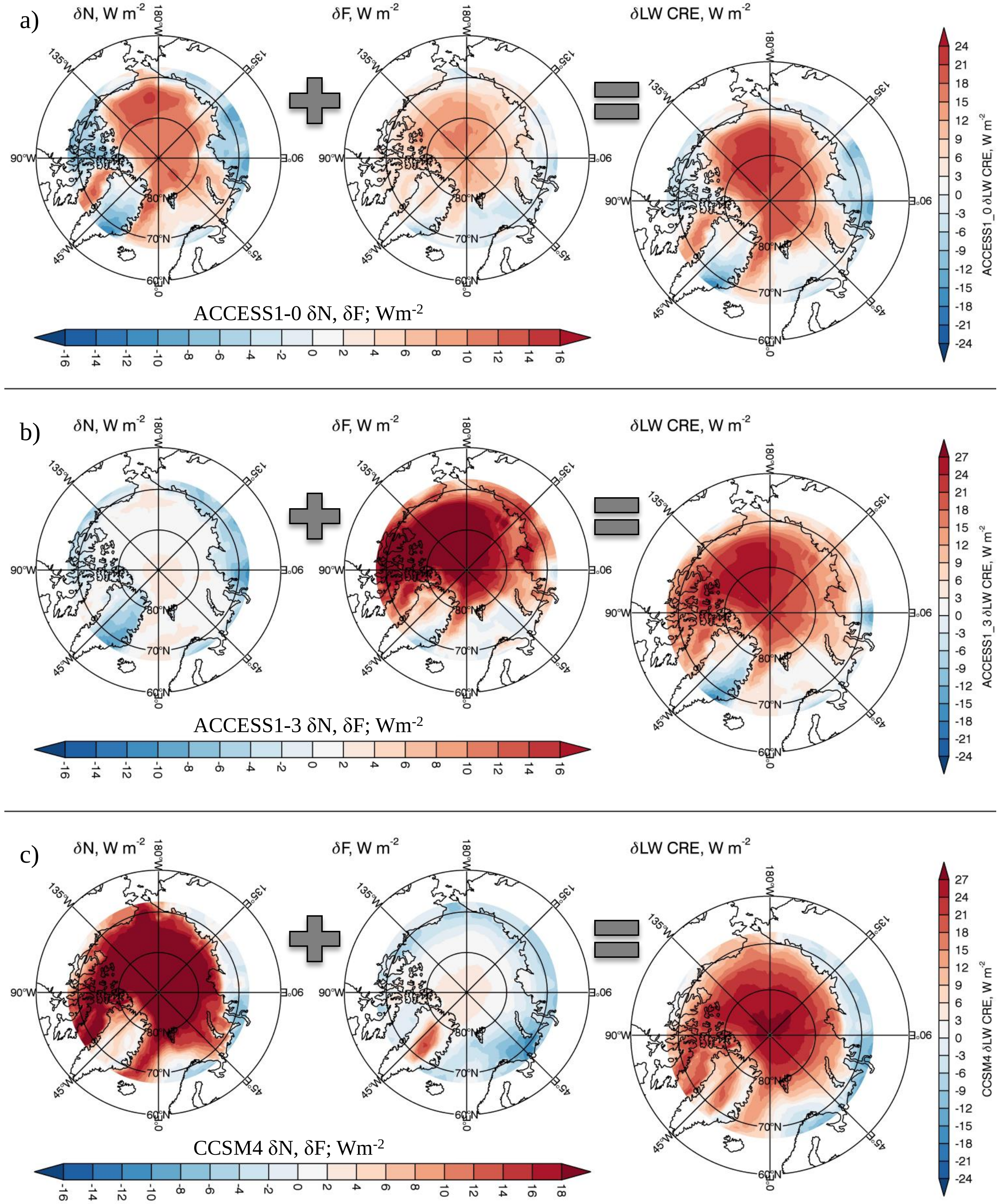

Figure 10. Spatial distributions of the LW CRE decomposition terms for (a) ACCESS1.0, (b) ACCESS1.3, and (c) CCSM4.

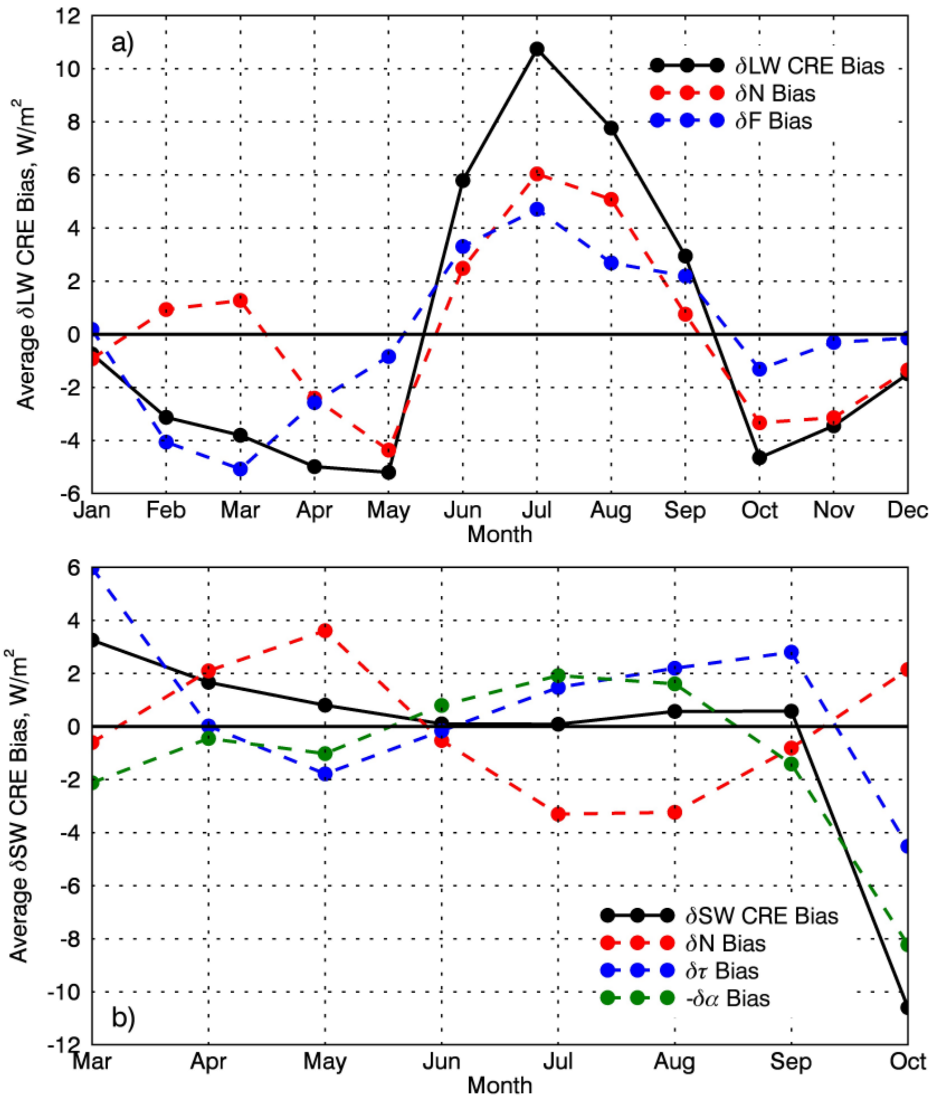

Figure 11. Arctic domain-average biases (Ensemble mean minus observations) for (a) LW CRE decomposition terms, and (b) SW CRE decomposition terms.

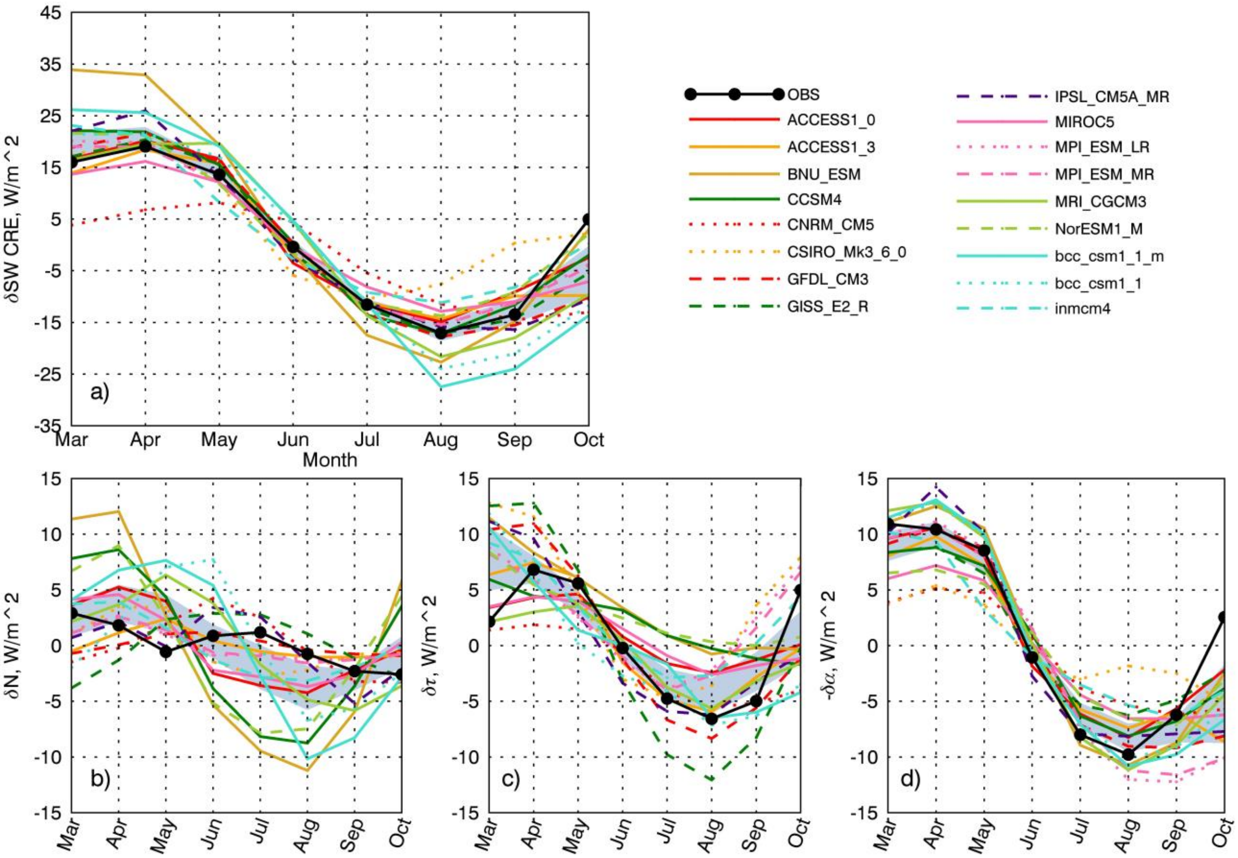

Figure 12. Arctic domain-average seasonal cycle of the SW CRE decomposition terms (a) δSW CRE, (b) δN, (c) δτ and d) -δα. Units are W m^{-2}.

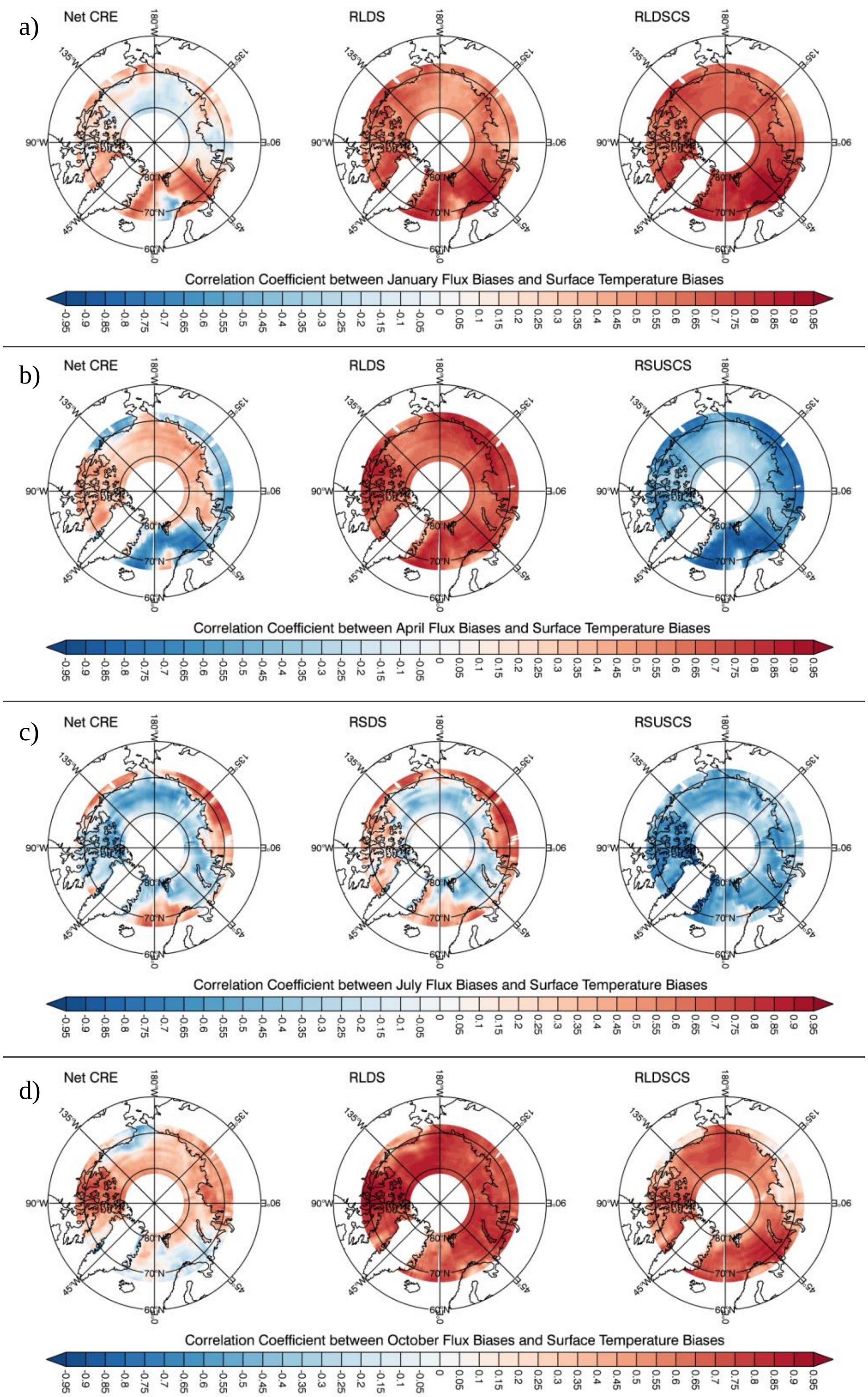

Figure 13. Spatial distributions of correlation coefficients between surface radiative flux biases and surface temperature biases for (a) January, (b) April, (c) July, and (d) October.